DOCUMENT RESUME

ED 263 742 FL 015 231

AUTHOR Derbyshire, Desmond C., Ed.
TITLE Work Papers of the Summer Institute of Linguistics, University of North Dakota Session, (Grand Forks, North Dakota 1985). Volume 29.
INSTITUTION North Dakota Univ., Grand Forks.
PUB DATE 85
NOTE 347p.
PUB TYPE Reports - Research/Technical (143) -- Collected Works - Conference Proceedings (021)

EDRS PRICE MF01/PC14 Plus Postage.
DESCRIPTORS *Acoustic Phonetics; African Languages; American Indian Languages; *Discourse Analysis; Distinctive Features (Language); French; Language Research; Linguistic Theory; *Morphology (Languages); *Phonology; Regional Dialects; Semantics; *Sentence Structure; *Sociolinguistics; Surveys; Syllables; Verbs
IDENTIFIERS Karitiana; Koh; Minnesota (Red Lake Falls); Mixe; Mixtec; Referents (Linguistics); Yagua; Zapotec

ABSTRACT

Preliminary versions of the papers from the 1985 Summer Institute of Linguistics presented at the University of North Dakota session include: "Referential Distance and Discourse Structure in Yagua" (Thomas E. Payne); "A Note on Ergativity, S', and S'' in Karitiana" (Daniel Everett); "Some Aspects of Zapotecan Clausal Syntax" (Stephen A. Marlett); "Syllable Structure and Aspect Morphology in Isthmus Zapotec" (Stephen A. Marlett and Velma B. Pickett); "Numi Mixtec Syllable Structure and Morphology" (Laura Gittlen and Stephen A. Marlett); "Fortis/Lenis Consonants in Guichicovi Mixe: A Preliminary Acoustic Study" (J. Albert Bickford); "The Inflectional/Derivational Distinction" (David Tuggy); "The Koh Verbal System" (Suellyn H. Glidden); and "Red Lake Falls, Minnesota: A Sociolinguistic Survey" (James Kapper). (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED263742

1985

WORK PAPERS

of the

Summer Institute of Linguistics

University of North Dakota
Session

Volume 29

Editor: Desmond C. Derbyshire

Summer Institute of Linguistics

FL015231

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Desmond C. Derbyshire

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

These are working papers and should not be cited without referring to their preliminary nature.

Orders and correspondence should be sent to:

SIL/UND Work Papers
c/o International Linguistic Center-Bookroom
7500 West Camp Wisdom Road
Dallas, TX 75236

3

PREFACE

Most of the papers in this volume present some provisional results of research being undertaken by members of the Summer Institute of Linguistics who are associated with SIL's summer program at the University of North Dakota, either as teachers or research students. The papers by Sue Glidden and Jim Kapper are slightly revised versions of their M.A. theses, which were approved by the UND Graduate School in 1984.

My thanks go to David Fortune and Pat Aldrich for handling the composition and production of the volume, to Susan Peterson and Betty Brown for proofreading, and to Hu Matthews and Tom Payne for their editing help.

D.C.D.

TABLE OF CONTENTS

Preface iii

Referential Distance and Discourse Structure in Yagua
Thomas E. Payne 1

A Note on Ergativity, S´, and S´´ in Karitiana
Daniel Everett 69

Some Aspects of Zapotecan Clausal Syntax
Stephen A. Marlett 83

Syllable Structure and Aspect Morphology in Isthmus Zapotec
Stephen A. Marlett and Velma B. Pickett 153

Ñumí Mixtec Syllable Structure and Morphology
Laura Gittlen and Stephen A. Marlett 175

Fortis/Lenis Consonants in Guichicovi Mixe: A Prelilminary Acoustic Study
J. Albert Bickford 195

The Inflectional/Derivational Distinction
David Tuggy 209

The Koh Verbal System
Suellyn H. Glidden 223

Red Lake Falls, Minnesota: A Sociolinguistsic Survey
James Kapper 283

REFERENTIAL DISTANCE AND DISCOURSE STRUCTURE IN YAGUA

Thomas E. Payne

1 Introduction
2 Coding devices
    2.1 A and S_a coding
        2.1.1 Verb coding (VC)
        2.1.2 Pre-predicate noun phrase (PNP)
        2.1.3 Verb coding plus (post-predicate) noun phrase (VC+NP)
    2.2 O and S_o coding
        2.2.1 Enclitic (E)
        2.2.2 Post-predicate noun phrase (post- NP)
        2.2.3 Enclitic plus noun phrase (E+NP)
    2.3 Oblique coding
        2.3.1 Head coding
        2.3.2 Noun phrase (NP)
        2.3.3 Head coding plus noun phrase (HC+NP)
    2.4 Summary
3 Topic continuity
    3.1 The contribution of topic continuity to a theory of anaphora
        3.1.1 Participant topicworthiness
        3.1.2 Discourse structure
    3.2 Technique
    3.3 Modifications
    3.4 Data base of the quantitative study
    3.5 The results
        3.5.1 A and S_a participants (subjects)
        3.5.2 O and S_o participants
        3.5.3 Oblique participants
        3.5.4 Summary of mean values by coding device
    3.6 Discussion
4 Exceptions to the general pattern
    4.1 Short-range devices used to code distant participants
        4.1.1 Contextual inference
        4.1.2 Levels of topicality
    4.2 Long-range devices used to code recently mentioned participants
        4.2.1 Ambiguity
        4.2.2 Elaboration
        4.2.3 Discourse promotion
        4.2.4 Dative objects
        4.2.5 Thematic structure
        4.2.6 Crucial inanimate participants
        4.2.7 After mention in quotes
        4.2.8 Residue
5 Conclusion
Appendix: The Non-identical Twins Cycle
Notes
References

1 Introduction

This paper concerns Yagua, a lowland language of Peru.[1] It also concerns the choices speakers make as to how they will refer to or ´code´ participants in discourse. The body of this paper will be organized into two broad parts corresponding to these two most general concerns. In the first part, consisting of Sect. 2, I will describe the major formal devices used to code participants in Yagua. In the second part, Sect. 3 and 4, I will look at the use of those devices in a body of folkloric narrative texts. I will take as starting point for the textual study the framework and methodology developed by Givón (1983a, b, c, d) for measuring ´referential distance´, i.e., the distance in number of clauses between one mention of a participant and its previous mention in the text. The general framework is outlined in some depth and several substantive modifications to the methodology are proposed. Deviations from the general predictions of Givón´s framework will then be examined in detail in Sect. 4. It is found that there are two circumstances under which relatively ´new´ participants can be coded with attenuated coding devices (i.e., verb coding and enclitics), and seven circumstances under which relatively recently mentioned participants can be coded with full noun phrases. Each of these circumstances represents an area of further research into the functional factors underlying the choice of participant coding devices in discourse.

2 Coding devices

The term ´coding devices´ (sometimes ´participant coding devices,´ Givón 1983a) will be used to refer to formal devices used to code (i.e. ´mention´ or ´refer to´) participants (as opposed to, e.g., actions). For Yagua these devices include noun phrases, pronouns, verb coding (agreement), enclitics and all combinations and ordering permutations thereof. In this section I will introduce and illustrate the major coding devices available in Yagua.

As a descriptive convention, I will follow Silverstein (1976) and Dixon (1979) in using the terms A, S, and O to refer to subjects of transitive verbs, subjects of intransitive verbs, and objects of transitive verbs, respectively. Furthermore, I will follow Dixon (1979) in using the terms S_a and S_o to refer to two classes of intransitive subjects. S_a participants are those intransitive subjects that are coded morphosyntactically like transitive subjects (i.e., S_a and A participants share the same set of possible coding devices), while S_o participants are those intransitive subjects that are treated morphosyntactically like transitive objects.

2.1 A and S_a coding

All A´s and S_a´s have the same set of possible coding devices in Yagua. These are verb coding (VC), pre-predicate noun phrase (PNP), verb coding plus post-predicate noun phrase (VC+NP), zero, pre-predicate pronoun, right-dislocated pronoun, and right-dislocated NP. In this paper I will be concerned exclusively with the first three of these devices, as the others are quite rare in discourse, and space dictates that I concentrate on the major coding devices of the language. (See T.

Payne 1985 for a more detailed description and functional analysis of coding devices in Yagua discourse.)

2.1.1 Verb coding (VC). A and S_a participants may be coded by verbal prefixes. The basic forms of the verb coding prefixes are identical to the head coding prefixes used on nouns to code possessors and on postpositions to code oblique objects. These basic forms are illustrated in Chart A:

	1st excl.	1st incl.	2nd	3rd	4th	inan
Sg.	ray-		jiy-	sa-		
Dl.	nááy-	vųųy-	są́ą́da-	naada-	jíy- (no number distinction)	rá-
Pl.	núúy-	vųųy-	jiryey-	riy-		

Chart A: Prefixal Forms[1]

T. Payne (1983b) presents rules that derive essentially all surface forms of person and number prefixes from the above underlying forms. There are four major classes of stems distinguished by the form of the initial syllable. Class I stems are all those whose initial syllable is a consonant other than **j** plus any vowel, or **j** plus the vowel **o.** The three other classes exhibit the initial stem syllables, **ja, ji,** and **ju** respectively. No stems begin with the syllable **je** in their underlying form, and **jo** stems conjugate just like Class I stems.

The following examples illustrate the VC device with a verb stem from each class:

(1) **Rañikyéé.**
ray-nikyee (Class I)
1SG-speak
´I speak.´ (S_a)

(2) **Sąątu** buyą́ą́.
sa-jatu buyąą (Class II)
3SG-drink manioc:beer
´S/he drinks manioc beer.´ (A)

(3) **Vuryiimíy.**
vurya-jimyiy[2] (Class III)
1PLINC-eat
´We eat.´ (S_a)

(4) **Sųųtúra.**
sa-jųtú-rà (Class IV)
3SG-carry-INAN
´S/he carries it.´

These prefixes may also occur on an auxiliary verb:

(5) **Vuryąą** jatú buyą́ą́.
vurya-ą jatu buyąą
1PLINC-AUX:IRR drink manioc:beer
'We will drink (or let's drink) manioc beer!'

(6) **Saniy** jiryiy yímújųra.
sa-niy jiryiy yí-mu-jų̀-rà
3SG-AUX:MALF grab 4-LOC-towards-INAN
'He grabbed it to himself (to his own detriment, or with evil intentions).'

2.1.2 Pre-predicate noun phrase (PNP). When a full noun phrase is used to code an A or S_a participant, that noun phrase may precede the verb. Verb coding then may not be used to code the same participant, as illustrated by the ill-formed strings following examples(7) and (8):

(7) **Sa-munaa-dee** kúútya.
3SG-placenta-DIM whisper
'His placenta^3 whispers.' (S_a)

*Samunaadéé sakúútya.
'His placenta he whispers.'

(8) **Jíryoonú** sų́ų́yanníí.
jíryoonú sų́ų́y-janu-níí
bushmaster bite-PAST3-3SG
'A bushmaster bit him.' (A)

*Jíryoonú sasų́ų́yanníí.
'A bushmaster he bit him.'

2.1.3 Verb coding plus (post-predicate) noun phrase (VC +NP). When verb coding is accompanied by the overt expression of a coreferential noun phrase within the clause, that noun phrase must follow the verb:

(9) **Sa-suvų́ų́y** Anita
3SG-afraid Anita
'Anita is afraid.' (S_a)

(10) **Sa-nááyi** Alchíco-rà
3SG-press Alchico-INAN
'Alchico presses it.' (A)

Conversely, when a full NP coding an A^4 or S_a participant follows the verb, a verb coding prefix is obligatory. Thus it is not necessary to annotate the coding device here termed VC+NP as to whether the NP occurs post-verbally or pre-verbally. It is a syntactic requirement that when an NP alone codes an A or S_a participant, that NP must be preverbal. When the NP occurs in combination with a coreferential VC prefix, however, then the NP must be post-verbal.

A preliminary hypothesis regarding the use of the post-verbal NP in addition to the VC prefix would be that the NP is uttered as an

´afterthought´ when the speaker judges that VC was not sufficient to uniquely distinguish the correct participant. However, there are four arguments against this analysis: 1) An inspection of texts reveals that the VC+NP forms are used more commonly and in less marked discourse contexts than the simple NP constructions. This would hardly be an expected characteristic of ´afterthought´ constructions. 2) VC+NP constructions are normally uttered under a single intonation contour, suggesting that they code a single focus of consciousness and therefore do not involve a reconsideration and partial restatement. 3) As illustrated in example (10), when an O participant occurs in the clause, it must follow any overt subject NP. This indicates that the overt subject NP is at least as closely tied syntactically to the verb as the O participant is, which again is not a reasonable characteristic of afterthought subjects. 4) Finally, there is another device--right-dislocated NP--that does, in fact, code afterthoughts, as evidenced by the fact that it is normally uttered under a new intonation contour and occurs very rarely in discourse. Thus it appears that the afterthought hypothesis does not explain the function of the VC+NP constructions, at least as far as the synchronic grammar is concerned.

2.2 O and S_o coding

As discussed in Sect. 1, the term O refers to participants which can generally be thought of as transitive objects. In Yagua, subjects of non-verbal predicates, e.g., predicate nominals, predicate locatives, etc., as well as certain intransitive verbal predicates, are treated morphosyntactically just like transitive objects. Such predicates have been termed S_o predicates following Dixon (1979), and the subject of such predicates is termed the S_o. In Yagua, the devices which code O and S_o participants are enclitic (E), post-predicate NP (post-NP), enclitic plus NP (E+NP), pre-predicate noun phrase, zero, pre-predicate pronoun, pre-predicate noun phrase plus enclitic, and right-dislocated noun phrase. Again, in this paper we will only be concerned with the three major coding devices, E, post-NP, and E+NP.

2.2.1 Enclitic (E). O and S_o participants in Yagua may be coded with an enclitic attached to the last post-verbal constituent. If no constituents (other than an NP coding the O or S_o itself) occur after the verb, the enclitic occurs on the verb itself. If a coreferential full NP also occurs, it immediately follows the enclitic. Chart B illustrates the paradigm for O and S_o enclitics.

	1st		2nd	3rd	refl	inan
	ex	inc				
Sg.	-ráy		-jíy	-níí		
Dl.	-nááy	-vų́ų́y	-sąądá	-naadá	-yù (no number distinction)	-rà
Pl.	-núúy	-vų́ų́y	-jiryéy	-ríy		

Chart B: Forms of O and S_o Enclitics

Examples (11), (12) and (13) illustrate the enclitic device used to code O participants of verbal predicates:

(11) Ravyą́ą́ta-**jíy.**
1SG:want-2SG
'I like/love you.'

(12) Sa-nááyi Alchíco-**ra.**
3SG-press Alchico-INAN
'Alchico is pressing it.'

(13) Sa-jáátya sínu-mu-**níí.**
3SG-toss land-LOC-3COL
'He tosses them on the shore' (i.e., fish, hence
the animate '3COL' enclitic).

As mentioned above, subjects of non-verbal predicates are treated morphosyntactically just like direct objects. Hence, they also can be coded with an enclitic. The following examples illustrate some simple predicate nominals in which the subject is coded with the enclitic device:

(14) a. Maésturu-**níí.**
teacher-3SG
'He is a teacher.'

b. Maésturu-**ríy.**
teacher-3PL
'They all are teachers.'

Examples (14a) and (14b) illustrate the simplest kind of predicate nominal construction. In these examples the predicate nominal is **maésturu**, and the subject is coded with an enclitic. In the following example the 'copula' **day** appears. This particle may or may not be classifiable as the same **day** that functions as a discourse particle. It is certainly not a verb since it has none of the properties characteristic of true verbs. However, it does frequently occur in predicate nominal constructions such as (15), and so I have glossed it 'copula' following Powlison (1969) in such contexts.

(15) Tomáása báárya dá**rya.**
Tomáása báárya day-rà
Tom thing COP-INAN
'It is Tom's thing.' (It belongs to Tom.)

In example (15) the form **day** is fairly strongly demanded, though in isolation the sentence is acceptable to native speakers without the **day.** In (14a) and (14b), **day** could occur immediately following the predicate nominal with no change in true functional meaning. A detailed study of the use of **day** from a discourse perspective is certainly an important area for future study.

Examples (16) and (17) illustrate the use of enclitics to code the subject of predicate locative clauses:

(16) Vóóka-ncha-**níí.**
cow-upon-3SG
'He is on the cow.'

(17) Rá-tuunu-dee-numaa-téé-**naadá.**
INAN-beside-little-now-INTS-3DL
'They (2) are right beside it now.'

In example (16) the predicate locative is **vóókancha** 'on the cow', and the subject is coded with the 3SG enclitic **níí.** In (17) the predicate locative is **rátuunudee** 'right beside it,' and the subject is coded with the 3DL enclitic **-naadá.** The formative **numáá** is a second position special clitic that modifies the sense of the predication. **Numáá** and several other second position clitics are also used in verbal predicates (see Doris Payne 1985 for a detailed discussion of clitic placement in Yagua).

In addition to non-verbal predicates, many intransitive verbal predicates in Yagua can appear with S_o coding, though S_a coding is always allowed as well. For example:

(18) a. **Naada**siimyaasíy rumusíy. (S_a)
naada-siiy-maasiy rumu-siy
3DL-run-exit there-from
'They (two) rush out from there.'

b. Rumusíy siimyaasi**ñaadá.** (S_o)
rumu-siy siiy-maasiy-naada
there-from run-exit-3DL
'From there they (two) rush out.'
or 'Out from there rush the two of them.'

In example (18a) the compound verb stem **siimyaasíy** occurs with the VC prefix **naada-**, while in (18b) the same stem occurs with an enclitic. Contrary to observations concerning S_o coding on verbs in other languages, there is no necessary semantic difference between (18a) and (18b) such that (18a) implies volition and control on the part of the subject whereas (18b) does not. Both of these sentences clearly imply volition, action, and control.

When verbs occur with S_o coding, they are formally parallel to non-verbal predications in that they may not indicate time reference. Certain time reference suffixes may appear on S_o predicates, but the meaning of these suffixes is then no longer that of time reference. For example, consider the following:

(19) Nuuñíítǫǫ**jásiy** musajomú.
nuudya-jítǫǫ-jásiy musa-jo-mu
1PLEX-arrive-PROX1 descend-NOM-LOC
'We arrived (earlier today) at the port.'

(20) **Siijééñuveeją́ą́siy** Mocáyu sų́tajyariñaadá.
siiy-jááy-nuvee-ja-jásiy Mocáyu sų́tay-jariy-naada
run-enter-ARR2-o'land-PROX1 Mocáyu shelter-into-3DL
'They two run overland there (close by) into Mocáyu's
shelter.'

In example (19) the PROX1 suffix **jásiy** imparts the time reference of 'earlier in the day of speaking' to the verbal predicate. Notice that this is a single participant clause with the VC prefix indicating the S_a participant 'we.' Example (20), however, is an S_o predicate as evidenced by the fact that the single participant 'they two' is coded with the enclitic **naadá**. The PROX1 suffix in (20) does not impart time reference, but rather specifies that the location of Mocáyu's shelter was near to the place where the twins exit from. This is completely consistent with the use of **jásiy** (and many other suffixes) with predicate nominals and predicate locatives. For example:

(21) Tąąri-ñiy jásiy sa-tuunu-naadá.
long:while-NIY PROX1 3SG-side-3DL
'They two are there at his side a long while.'

Not: 'They two were (earlier today) at his side
a long while.'

(22) Tą́ą́ra sííva jásiy?
tą́ą́ra sa-íva jásiy
what 3SG-DAT PROX1
'What does he have there?'
(Lit: 'What is to him there?')

Not: 'What did he have (earlier today)?'

(23) Ráju-ra mákindya-jásiy.
many-INAN machine-PROX1
'Lots of machines are there.'

Not: 'Lots of machines were there (earlier today).'

These formal facts illustrate that verbal predicates with S_o coding have much in common with predicate nominals. However, on semantic grounds we must say that such predicates are not nominals. Specifically, there is no sense in which example (18b) can mean, 'There are the rushed out ones.' The stem **siimyaasíy** can only be interpreted as a verb in that 1) it cannot fill the role of a noun phrase, e.g., subject or object, in a clause, and 2) it must take nominalizing suffixes in order to fill such a role. This is true for all verbs that can occur with S_o coding.

The use of a locative demonstrative (**rumusíy** in example (18b) above) correlates highly, but not absolutely, with the presence of S_o coding. Example (24) illustrates an S_o verb with the locative demonstrative **mííy**:

(24) Múúy kinchunuveejásiryíy.
múúy kinchu-nuvee-jásiy-ríy
there light:fire-ARR2-PROX1-3PL
'There upon arrival they light a fire.' (KT57)

Occasionally, however, verbs occur with S_o coding without a preceding locative demonstrative:

(25) Siiryɨ̨ɨ̨dañíí koodidyéé.
siiy-rɨ̨ɨ̨-day-níí koodiy-dee
run-in:passing-immediately-3COL snake-DIM
'Immediately the little snakes[5] scurried.'

(26) Kuutya-nuvaa-níí,
whisper-ARR2-3SG
'He whispers on arrival, . . .'

In summary, we have seen that enclitics are used to code three general classes of participants: 1) objects of transitive verbs, 2)subjects of non-verbal predications such as predicate nominals and predicate locatives, and 3) subjects of certain intransitive predicates based on verbal stems but sharing many features of non-verbal predicates.

2.2.2 Post-predicate noun phrase (post-NP). A full post-predicate NP may code an O or S_o participant, in which case an enclitic is 'optional' from a sentence grammar perspective. The following examples illustrate the simple post-predicate NP device with no coreferential enclitic:

(27) Sąątú **buyą́ą́.**
sa-jatu buyąą
3SG-drink manioc:beer
'S/he drinks manioc beer.' (O)

(28) Jiryimimyáá **raryą́ą́vątá.**
jiy-rimiy-maa ray-rą́ą́vą-ta
2SG-spill-PERF 1SG-poison-partitive
'You spilled part of my poison!' (O)

2.2.3 Enclitic plus noun phrase (E+NP). However, the enclitic can co-occur with a coreferential NP. This I will term the E+NP device:

(29) Sa-suutá-**ra** **sújay.**
3SG-wash-INAN clothes
'S/he washes the clothes.' (O)

(30) Maesturu-**níí Alcídes.**
teacher-3SG Alcides
'Alcides is a teacher.' (S_o)

When the E+NP device is used, the NP always immediately follows the enclitic, and no constituent may intervene between the enclitic and the coreferential NP. These facts constitute evidence that the enclitic forms a constituent with the following noun phrase, even though it is phonologically attached (cliticized) to whatever word precedes. It is suggested in T. Payne (1983a) that the difference between the E+NP and NP constructions has to do with definiteness--the enclitic is used for definite O´s and is not used for indefinite O´s. However, there are clear cases where definite O´s are coded without the enclitic and where indefinite O´s are coded with the enclitic. For example, possessed O´s occur with or without the enclitic in about the same proportions. This fact would not be expected if the presence of the enclitic were strictly dependent on definiteness of the O, since possessed NPs are overwhelmingly definite (see, e.g., DuBois, 1980:208). Example (31a) illustrates the use of a possessed O nominal without a preceding enclitic, and (31b) illustrates a possessed S_o nominal without the enclitic:

(31) a. Saryey jɨɨta**-0 jíchikidií.**
sa-ryey jɨɨ́ta-0 jíy-sikidii
3SG-grab JIITA-0 4-intestines
´He grabs his own intestines.´ (TC519)

b. Múúñumaatéé**-0** **rájaachéy.**
múúy-numaa-téé-0 rá-jaay-say
there-now-INTS-0 INAN-heart-CL:tree
´There now is its (the tree´s) heart.´ (TC327)

Example (32) illustrates that the E+NP device can be used to code indefinite participants. In this example the coca leaves have not been previously mentioned in the discourse, and the likelihood that they have been implicitly mentioned or are perpetually identifiable is minimal. Therefore we conclude that the leaves are indefinite in the sense of DuBois (1980), Chafe (1976), and others. And yet in this sentence the enclitic is used:

(32) Saruuy**ę́ę́rya** **japatíy.**
sa-ruuy-yęęy-rà japatiy
3SG-roast-DIST- INAN coca
´He is roasting coca.´ (TC536)

In T. Payne (1985 chapter 4) a more satisfactory explanation for the use of the E+NP device is presented. This explanation rests on the notion of ´referentiality´ (DuBois 1980), persistence (Givón 1983a, b, c), ´discourse manipulability´ (Hopper and Thompson 1984), or ´deployability´ (Jaggar 1985). Briefly, the simple post-predicate NP device is used for O and S_o participants that do not persist on the discourse stage, i.e., they are not ´destined´, as it were, to figure prominently in the immediately ensuing discourse. The E+NP device, on the other hand, is used for participants that will persist for a greater span of text. This explanation is quantitatively justified for Yagua in T. Payne (1985).

2.3 Oblique coding

Obliques are coded with exactly the same set of coding devices as are possessors: HC, NP, and HC+NP. The head of an oblique phrase is the postposition indicating the semantic role of the oblique nominal. In the following subsections, examples of each of the coding devices used for postpositions are illustrated.

2.3.1 Head coding (HC). Postpositions and nouns fall into the same conjugation classes as do verbs. Head coding prefixes referring to the oblique participant whose semantic role is indicated by the postposition are the same as those referring to subjects (see chart A). The following examples illustrate some simple postpositional phrases where the participant is coded with an HC prefix on the postposition:

(33) **Rá**-tąąsá
INAN-middle
´In the middle of it´ (Class I)

(34) Sąąríy
sa-jaríy
3SG-underneath
´Underneath him´ (Class II)

(35) Sųųsíy
sa-jųsíy
3SG-from/after
´From/after him´ (Class IV)

2.3.2 Noun phrase. Like subjects with respect to their verbs, oblique nouns can occur as full NPs immediately preceding their postpositions. In such a case an HC prefix may not occur on the postposition:

(36) **Núú** tąąsá
path middle
´in the middle of the path´

(37) Tą́ą́ra **rudamú** sayasíy?
tą́ą́ra ruda-mu sa-jiya-siy
what day-LOC 3SG-go-PAST1
´On what day did he go?´

2.3.3 Head coding plus noun phrase (HC+NP). When HC does occur on a postposition, a full NP referring to the oblique nominal may follow:

(38) **Rá**-tąąsa **sa-moo-mu-dáy**
INAN-middle 3SG-face-LOC-DAY
´Right on his face´

(39) Riñeechǫ́ **munuñúmiy**
riy-naachǫ munuñu-miy
3PL-towards savage-PL
´Towards the savages´

Postpositional phrases are isomorphic with possessive constructions. In some cases stems which are clearly nouns are used in postpositional phrases to modify the sense of the relation being expressed. For example, the stem **moo** ´face/forehead´ illustrated in example (38) is used in combination with postpositions to impart the sense of 'in front of.' For example, contrast the a and b examples below:

(40) a. Sa-moo-mú
3SG-face-LOC
'In front of him' (Lit.: 'at his fact')

b. Siimú
sa-jimú
3SG-LOC
'On him' or 'at his house'

(41) a. Sa-moo-mu-síy
3SG-face-LOC-from
'Away from where he is headed' (Lit.: 'away from his face')
b. Siimusíy
sa-jimu-siy
3SG-LOC-from
'Away from him' (implies he is not moving)

2.4 Summary

In this section I have illustrated the major devices used to code participants in Yagua. A and S_a participants employ the devices I have termed VC, PNP, and VC+NP, while O and S_o participants employ the devices E, NP, and E+NP. Oblique participants employ the devices HC, NP, and HC+NP. Though I have termed the verbal prefixes VC and the postpositional (and possessor) prefixes HC, it is understood that the prefixes involved in both of these devices are selected from the same set.

3 Topic continuity

In this section I pose the question: How do discourse pragmatic factors affect the choice of participant coding devices in Yagua discourse? In attempting to answer this question I tentatively adopt the framework set out by Givón (1983a, b, c) for assigning quantitative continuity values to the various coding devices. This approach is a logical first step in that it rests on a well-defined limiting hypothesis, namely, that the use of all coding devices can be explained in terms of 'continuity' (as defined below). Once this preliminary hypothesis is tested, deviations from the expected results will point out other factors affecting choice of coding devices, thereby defining areas for further investigation. In Sect. 4 I examine specific examples that violate the general findings of the topic continuity study, and attempt to determine what additional factors might be influencing the choice of coding devices.

3.1 The contribution of topic continuity to a theory of anaphora

According to Givón (1983b:7), 'The **clause** ('sentence') is the basic information processing unit in human discourse.' Discourse is made up of chains of clauses, which are in turn combined into larger units called 'paragraphs,' 'sections,' 'chapters,' etc. In order to qualify as a discourse, a chain of clauses must 'hang together' in certain demonstrable ways, i.e., it must be **continuous.** This particular observation certainly does not originate with Givón, but plays a central role in much previous work on discourse structure, though under different terminology. For example, Halliday (1967), Halliday and Hasan (1976), and Grimes (1975), among others, use the term 'cohesion' for essentially this same concept. Givón's central insight is that continuity (cohesion, if you will) from one clause to the next in real discourse is the most expected, unsurprising, and unmarked situation. Discontinuity is unexpected, surprising and marked. This observation suggests an 'iconicity principle' of human communication in general, and of topic continuity in particular:

(42) 'The more disruptive, surprising, discontinuous or hard to process a topic is, the more **coding material** must be assigned to it.'

(Givón 1983b:18, emphasis in the original.)

This principle is consistent with a more general and obvious behavioral principle: 'Expend only as much energy on a task as is required for its performance' (ibid.). Principle (42) is iconic in that it relates a formal coding scale (amount of coding material) to a scalar functional domain (continuity) in a non-arbitrary way, i.e., the less continuity the more coding material. Characteristically, then, continuity (of various sorts) is either not marked morphologically or is encoded with minimal morphological marking. Discontinuity, on the other hand, is encoded with more substantial morphological marking, or with otherwise more 'marked' morphosyntactic structures.

There are three types of continuity mentioned in Givón (1983b): thematic continuity, action continuity, and participant, (or topic) continuity. Although he acknowledges that these three kinds of continuity are intimately related to one another, Givón concentrates on topic continuity for purposes of the quantitative analytical procedure he proposes. In Givón's view 'topic' is seen as a scalar category. Participants are more or less topical at any given point in a discourse. It is more continuous (and therefore less surprising, or less marked) for a speaker to refer to (or mention) a more highly topical participant than a less highly topical one.[6] Certainly there are times when a speaker needs to refer to participants that are low in topicality. In particular, participants that have not been brought onto the discourse stage have no topicality whatsoever.[7] Therefore, in order to introduce a participant, or to reintroduce a participant after a significant period of absence, more marked morphosyntactic coding devices are called for. The speaker/writer must 'work harder' to signal discontinuity since it is not the most natural state of affairs in human discourse. Givón's quantitative method is a way of determining how topical any participant is at any given point in a text. Once topicality is determined in a rigorous, non-circular, non-impressionistic way, participant coding devices can be ranked in terms of the average topicality values of the participants they code.

Certainly an all-inclusive theory of anaphora would have to take into account many factors other than brute number of mentions and number of clauses since previous mention in order to accurately characterize the use of participant coding devices in discourse. These other factors fall into two categories: 1) factors involving the **topicworthiness** of the participants themselves, and 2) factors involving the structure and flow of the discourse. These two factors will be discussed in the following two subsections, respectively.

3.1.1 Participant topicworthiness. Participant topicworthiness is here defined as the relative likelihood for a participant to be ´talked about´ (Reinhart 1982). There are at least two kinds of topicworthiness associated with any potential participant in a discourse: **inherent** topicworthiness and **context-imparted** topicworthiness. These notions are reminiscent of, but not identical to, DuBois´ notions of **intrinsic salience** and **plot salience** (DuBois, 1980:248-49). Certain entities are inherently more topicworthy than others, e.g. humans are more likely topics than non-humans, animates are more likely topics than inanimates, etc. Other entities are likely topics because of the semantics or pragmatics of the particular speech context. This second kind of topicworthiness is that which will be the focus of the rest of this paper and hence merits some elaboration here.

To take an obvious case of context-imparted topicworthiness, the speech act participants themselves are always highly topicworthy. They are always ´available´ for reference in any discourse since interlocutors must always be conscious of one another. This fact explains why speech act participants can be universally referred to with reduced morphosyntactic coding devices, such as first and second person pronouns, regardless of whether or how many times they´ve been mentioned previously in the discourse. In addition, however, particular speech situations may impart topicworthiness to certain classes of entities. For example, in certain religious circles God is perpetually a highly topicworthy entity. Hence, in the context of a religious gathering of this type, a participant coding device of the appropriate inflectional category but with no obvious antecedent is taken to be a reference to God. Another more commonplace example of this same phenomenon is the fact that entities can be non-linguistically or implicitly brought onto the discourse stage, by deixis, inference, body language and shared presuppositions of the speech act participants. Also, the choice of a particular verb causes the class of participants that are typical for that verb to be more topicworthy. For example, ´to spew´ in English implies a liquid participant; ´to speak´, a human participant, etc. Finally, in narrative the central characters of the story are more likely topics than the non-central characters. In short there are a myriad of factors that affect ´degree of topicworthiness,´ many of which are not accessible to the linguist working from transcribed materials, especially when the linguist does not have native understanding of the language or of the culture of the speakers of that language. Topicality (in the sense of Givón 1983b, as number or density of mentions) is only one kind of context-imparted topicworthiness. That is, if a participant has already been mentioned a lot and/or has recently been mentioned in the current discourse, it is more likely to continue as topic (it is more topicworthy) than are other entities, other factors being equal.

Despite the fact that the topic continuity methodology is sensitive to only a few of the factors involved in topicworthiness, it does represent a step in the right direction. The value of this methodology is that it is quantitative, rigorous, and non-impressionistic. Future research will certainly refine the methodology and the theoretical principles that underlie it until all variables have been isolated and incorporated into the findings.

In the quantitative study of Yagua presented in this paper, I have attempted to control for some of the variables outlined above by 1)distinguishing between central, major and minor characters (Sect. 3.3), 2) excluding from the general topic continuity counts references to entities which are clearly perpetually topicworthy due to their universal presence on the discourse stage, e.g., ´the sun,´ ´the day,´ etc. and 3)excluding references to speech act participants. The other factors (e.g., deictic mentions, implicit mention due to semantic subcategorization of verbs) will only be dealt with as they are obviously relevant in particular cases.

3.1.2 Discourse structure. The second major group of factors affecting the choice of coding devices has to do with the structuring of the information contained in a text. Again there are two subtypes of factors: 1) the hierarchical thematic or episodic structure of the text, and 2) factors of grounding, i.e., foregrounding and backgrounding of information (Hopper and Thompson 1980). The first factor corresponds to Givón´s ´thematic continuity´ and relates to the fact that thematic (or episode) boundaries crucially affect a speaker´s choice of coding devices. The second group of factors corresponds to Givón´s ´action continuity.´ Neither of these areas is dealt with in great detail in Givón (1983a, c, d), though he does compare continuity indices for participants at thematic junctures with those not at such junctures (1983d:192ff.).

Two recent works which incorporate the notion of thematic structure into the question of choice of participant coding devices are Clancy (1980) and Fox (1984). Clancy (1980) is the first work in which topicality (though Clancy does not use this term) is measured in terms of distance in number of clauses since last mention. However, in addition to counting numbers of clauses, Clancy also notices that discourse boundaries tend to elicit stronger coding devices than would be expected given a strictly linear view of continuity based on number of mentions or distance since last mention. The particular boundaries that Clancy finds relevant are ´world shifts,´ where a narrator shifts between the ´real world,´ i.e., the situation in which the narrative is being recounted, and the ´story world´, i.e. the world being depicted in the narrative. Clancy also finds ´episode´ boundaries significant. There are many other factors that Clancy deals with in her article on a case-by-case-basis.

Fox (1984) shows that the choice between use of a pronoun versus a full noun phrase in English is influenced by the hierarchical structure of the content of the text. The frameworks that Fox employs for determining the hierarchical structures of texts are Rhetorical Structure Analysis (Mann and Thompson 1983) for written texts, and Conversational Analysis (Sacks, Schegloff, and Jefferson 1974) for conversational texts. Fox takes particular examples of what would, from a strictly linear point of view based on number of clauses since previous mention, appear to be excessively strong coding devices (i.e., full NPs) or excessively weak devices (i.e., pronouns) and shows how from a hierarchical point of view, such patterning is explainable (Fox 1984:240ff.). On the basis of such examples Fox rejects the ´distance´ view of topicality in favor of a hierarchical view.

There is no question that hierarchical structure must be taken into account in an all-inclusive theory of anaphora. However, the essential contribution of Givón´s work on topic continuity is that it provides a quantitative method of calculating at least some of the factors affecting choice of coding devices. It is not meant to be all-inclusive or predictive, in the sense of being able to provide an algorithm for generating exactly the correct coding choices in a text, and none of the incorrect ones. It simply provides a rigorous, quantitative method of comparing the functions of coding devices according to certain well-defined parameters (referential distance, persistence and ambiguity as defined below). Future research on the use of anaphora in discourse must still provide a rigorous, preferably quantitative way of characterizing the effect of thematic continuity on participant coding choices. In Sect. 4.1.2 and 4.2.5 I will illustrate two respects in which the hierarchical structure of the text affects the choice of coding devices in Yagua.

3.2 Technique

In this paper topicality will be measured in terms of the index of **referential distance.** Two other indices, persistence (´decay´) and ambiguity (´interference´), will not be considered here for reasons of space limitations. T. Payne (1985) shows that the functional domains measured by the indices of persistence and ambiguity are sufficiently distinct from that measured by referential distance so as to render each worthy of separate treatment, though all fall into the overarching functional category of topic identification and manipulation in discourse.

The index of referential distance (RD) is based on the assumption that participants that have not been mentioned recently in the discourse are more difficult to process and are therefore less topical than those that have been mentioned more recently. This index measures the gap between the current mention of a participant and its previous mention in the discourse in terms of number of clauses. Thus an RD of 1 indicates that the participant was last mentioned in the immediately preceding clause and is therefore maximally continuous. In the extreme case of discontinuity, where a participant has not been mentioned at all in the present discourse, the RD index is technically infinite. However, since we cannot deal satisfactorily with infinite values, I will follow Givón (1983a, c, d) in imposing the arbitrary limit of 20 on the RD index. Thus participants which are introduced into the discourse for the first time, or are absent from the discourse stage for twenty clauses or longer receive the RD index of 20.

3.3 Modifications

The major modifications I will introduce into Givón's methodology involve the treatment of quoted material. This is particularly important for this study since over thirty percent of the clauses in the corpus are quotes. Participant mentions within quotes will not be assigned RD measurements, since I judge that their continuity pertains more to the quoted discourse than to the discourse in which the quote appears. However, such mentions will be counted for purposes of determining the RD of other mentions that occur outside the quotes. This decision rests on the assumption that if a participant is mentioned in a quote, that participant is 'on stage' and therefore is just as potentially topical as it would be were it mentioned outside a quote.[8]

A second modification that I will introduce will be to count the character that utters the quote as having been mentioned, even if no overt mention of that character is made. This modification is based on the assumption that if a participant has a 'speaking part' in the discourse drama, that participant must be 'on stage,' and therefore must be available as a discourse topic at the point where he or she speaks. It is quite common in Yagua discourse for a series of quotes to occur in which two or more participants are interacting, but where explicit verbs of saying are not indicated for every conversational 'turn' in the quoted discourse. It is simply understood by the content of the turn which interactant is speaking. Such situations will be considered to constitute a mention of the speaker for purposes of calculating the RD of other mentions. However, such implicit mentions will not themselves be assigned an RD measurement.

Other modifications involve the measurement of mentions of participants that are referentially included within mentions of other, non-singular participants. Non-singular mentions are considered to be mentions of each of the individuals contained in the group. Thus the RD of a non-singular mention would be the distance back to the last mention of any of the included individuals. Similarly, non-singular mentions are considered to be mentions of each of the individuals for purposes of measuring the RD of other singular mentions of those individuals.

Finally, I will introduce four **character statuses** as follows:

1. Central characters--These are the characters that the text is about, and which are normally present throughout the text. Central characters do not lose their status as central characters, even if they are not mentioned for an entire episode (e.g., the snail episode of the Twins Cycle, TC398 to TC458, see appendix).

2. Non-central major characters — those characters mentioned five or more times in twenty clauses in a single episode. Characters mentioned only four times in twenty clauses are also considered major if in the majority of those mentions (i.e., three or four) the character was a subject, i.e., an A or S participant. For purposes of determining whether a character is major or not, a quote is considered to be a (non-subject) mention of the character who utters the quote. Major characters may lose their status as major in a subsequent episode if they don't meet the criterion.

3. Perpetually present--This category encompasses that small number of referents that are automatically present on the discourse stage and therefore do not need to be introduced, e.g., the sun, the day, etc. Also, dummy referents such as the subject of ´thunder´ in, ´It thunders.´ Mentions of such participants are not given continuity indices, and therefore do not figure into the counts for the various coding devices.

4. Minor characters--All other participants.

The scope of this study is purposely limited in several respects. I will be concerned only with A, O, S (all types, including subjects of non-verbal predicates) and all Oblique (OBL) participants. I have not calculated continuity values for interjections or possessors, though these categories of mentions are considered in the measurement of RD for other mentions. Furthermore, as mentioned in Sect. 1, there are many coding devices in Yagua that I will not be considering, e.g., zero, pronouns, devices involving right or left dislocation, etc. These devices are all somewhat ´marginal´ numerically in Yagua discourse, and all have functions outside of the domain measured by the index of referential distance (T. Payne 1985; See Doris Payne 1985 for an explication of the pragmatically marked status of right and left dislocations).

3.4 Data base of the quantitative study

The data base for this study consists of four texts, all of which are essentially folkloric narratives. Table 1 summarizes the data base of this study:

Text	Number of Clauses[9]		
	Quotes	Non-Quotes	Total
1. First Squirrel (FSQ)	49	87	136
2. Hunter Narrative (HN)	72	154	226
3. Kneebite Twins (KT)	32	78	110
4. Twins Cycle (TC)	232	371	603
Totals:	385	690	1075

Table 1: Corpus for the Topic Continuity Study

The fact that all these texts are folkloric narratives biases the sample, to be sure. However, I have chosen to keep the genre constant in order to eliminate genre as a possible complicating factor as much as possible. Also, folkloric narrative is virtually the only non-first person genre available in Yagua, and questions of topic continuity are not as relevant when the primary characters are speech act participants for the simple reason that all speech act participants are always highly topicworthy. Finally, since little is known about Yagua in general, I judge it appropriate to begin the investigation with straightforward, narrative material. Future research would certainly need to include other genres.

These texts were all orally composed by unquestionably competent native storytellers, under reasonably natural circumstances. Transcriptions and translations were made either by native speakers, or by linguists working closely with the storytellers themselves (all of whom were preliterate at the time of storytelling). Text 1 appears in Appendix 2 of T. Payne (1985) and is an episode of a longer Yagua folktale titled ´Little Baldy´ in Powlison (1969). This particular version was recorded as a self-contained unit, so all continuity is relevant within the span of text examined. The other texts were all recorded by Paul Powlison. Texts 2 and 3 remain unpublished, while Text 4 appears in Powlison (1969), though in a different orthography than that which is used in this study, and without clause numbering. The free translation of Text 4 appears in the appendix of this paper.

3.5 The results

The tables and discussion presented in this section represent only a few of the possible ways of displaying the results of the quantitative study. I have endeavored to provide enough figures so that the reader with interest in a specific issue not dealt with in the discussion will be able to glean the appropriate data from the tables provided. Inevitably, however, the discussion is limited to a few areas of particular interest to this thesis. The following abbreviations will be used in the tables: C = Central character, MA = MAjor character, MI = MInor character (see Sect. 3.3 for an explanation of these terms), T = totals, n = number of instances, Mn = mean. A, O, S_a, and S_o will, of course, refer to semantico-syntactic roles, as outlined in chapter 1, Sect.1.5.2.

Mean RD values have been calculated and appear on the bottom row of each table. Since individual deviations from those mean values will be of particular interest, however, the tables also indicate the total number of instances of each value of each index for each coding device. Thus the tables graphically represent the distribution of instances of particular values within the entire range of possible values.

3.5.1 A and S_a participants (subjects). Table 2 indicates the referential distances for all A and S_a participants coded with the VC (verb coding) device:

RD	A				S_a				Total
	C	MA	MI	T	C	MA	MI	T	
1.	81	43	2	126	123	64	9	196	322
2.	13	12	1	26	22	23		45	71
3.	7	4		11	10	8		18	29
4.	2	1		3	6	3	1	10	13
5.	2			2	1	1		2	4
6.		1		1		1		1	2
7.									
8.									
9.						1		1	1
10.						3		3	3
.	(no instances of VC device coding RD of 11 to 19)								
.									
20.	1			1	3			3	4
n =	106	61	3	170	165	104	10	279	449
Mn=	1.57	1.46	1.33	1.52	1.62	1.88	1.3	1.77	1.68

Table 2: Referential Distance, Verb Coding Device
A and S_a Categories

From Table 2 we observe that the referential distance for the VC device is quite low and is roughly equal for all participant categories. The majority of instances of the VC device (322/448 or 72%) code participants which had been mentioned in the immediately prior clause (RD = 1), and most of the others cluster nicely near the lower end of the scale. These gross facts lead me to classify VC as a **short-range** coding device. That is, its primary function is to code participants that have been mentioned very recently in the discourse. However, it is interesting to note that occasionally the VC device is used to code quite distant participants. In fact, four times in this corpus the VC device has the maximal RD index of 20, even though it never has an RD index between 11 and 19. In Sect. 4 I will look at specific examples of VC used to code relatively distant participants (RD>8) for possible explanations for this patterning.

Table 3 presents the referential distance counts for the PNP device used to code A and S_a participants:

RD	A				S_a				
	C	MA	MI	T	C	MA	MI	T	Total
1.		2		2	1		1	2	4
2.									
3.					1			1	1
4.									
5.							1	1	1
6.						1		1	1
7.									
8.	1			1					1
.									
.									
20.			1	1		1	5	6	7
n =	1	2	1	4	2	2	7	11	15
Mn=	8	1	20	7.5	2	13	15.14	12.36	11.07

Table 3: Referential Distance, Pre-Predicate NP Device
A and S_a Categories

An obvious and expected finding of the figures in Table 3 is that PNP is a relatively **long-range** coding device. That is, it typically codes participants that have been absent from the discourse stage for a substantial amount of time. The mean RD of 11.07 for the PNP device is significantly higher than 1.68 for the VC device. However, the figures in Table 3 do not show nearly as much homogeneity as do those for the VC device. In particular, we notice that S_a participants taken as an aggregate exhibit higher referential distance than do A participants (RD = 12.36 for S_a and RD = 7.5 for A). This fact indicates that participants coded in the S_a role (subjects of intransitive verbal predicates) are generally less topical and more discontinuous than those coded in the A role (transitive subjects). This observation is consistent with DuBois (1981) who observes that intransitive subjects and transitive objects share the characteristic of being the primary roles in which ´new´ information is introduced, as opposed to a transitive subject which is typically ´given´ information. And in fact a quick glimpse ahead reveals that the pre-verbal NP device used to code O participants has an RD more similar to that of the same device used to code S_a rather than A participants (see Table 8).

The individual figures for central, major, and minor participants in Table 3 are not particularly helpful since the number of examples is so low. However, we may make two general observations: 1) the PNP device is relatively uncommon, occurring only fifteen times in 690 non-quote clauses of text, and 2) when it is used, it tends to code non-central characters (twelve instances) rather than central characters (three instances).

Table 4 presents the referential distance counts for the device termed verb coding plus (post-verbal) noun phrase (VC+NP) used to code A and S_a participants:

RD	A				S_a				
	C	MA	MI	T	C	MA	MI	T	Total
1.	2	6		8	10	11		21	29
2.	1	1		2	4	6	1	11	13
3.	1			1	3	1	3	7	8
4.						3		3	3
5.	1			1	1			1	2
6.		1		1	1	3		4	5
7.		1		1			1	1	2
8.									
9.					2			2	2
.									
.									
18.					1	1		2	2
19.									
20.	1	2	4	7	2	9	9	20	27
n =	6	11	4	21	24	34	14	72	93
Mn=	5.33	5.55	20	8.24	4.75	6.94	14.14	7.86	7.95

Table 4: Referential Distance, Verb Coding Plus NP Device
A and S_a Categories

The aggregate mean for the RD of the VC+NP device (7.95) is not as high as that of the PNP device (11.07). Thus I conclude that participants coded with the PNP device are **less topical** than those coded with the VC+NP device. This fact is consistent with the view expressed in Givón (1983c:19) that left-dislocated NPs are universally associated with higher discontinuity than are NPs in the neutral position.[10] However, these same facts seem to go against the scale of phonological size (Givón 1983c:18), which says that the phonologically ´larger´ device (in this case VC+NP) should be correlated with higher discontinuity than the ´smaller´ device (PNP). Clearly there is some other factor involved here. This issue is discussed briefly in T. Payne (1985) chapter 8.

Again, the figures in Table 4 exhibit less homogeneity than those for the VC device. Unlike the PNP device (Table 3), the referential distance counts for the VC+NP device do not separate out according to semantico-syntactic role, i.e., for the VC+NP device A and S_a participants as a whole exhibit similar RD counts (8.24 and 7.86, respectively). However, what we notice in Table 4 is that the counts do separate out according to character status. Central characters are consistently more topical, in terms of RD, than major characters, which are in turn consistently more topical than minor characters.

3.5.2 O and S_o participants. Table 5 indicates the Referential Distances of O and S_o participants coded with the simple E (enclitic) device. A distinction is also drawn between S_o participants of predicate nominals, and S_o participants of verbal predicates.

RD	O				S_o (pred nom)				S_o (verbal)				
	C	MA	MI	T	C	MA	MI	T	C	MA	MI	T	Total
1.	14	22	7	43	1	3	1	5	7	2		9	57
2.	5	12		17						3		3	20
3.	4	2	1	7					1			1	8
4.		2		2					1			1	3
5.	1	1		2									2
.													
.													
.													
20.		2	1	3									3
n =	24	41	9	74	1	3	1	5	9	5		14	93
Mn=	1.71	2.56	3.33	2.38	1	1	1	1	1.55	1.4		1.57	2.18

Table 5: Referential Distance, Enclitic Device
O and S_o Categories

From Table 5 we can conclude that, like the VC device, the enclitic is a short-range device in that most instances of this device cluster nicely towards the low end of the RD scale. Another interesting parallel between the enclitic device and VC is that for both there are some instances of the device used to code quite distant participants even though there is a significant gap in which no instances of the device are found. For the enclitic device the gap is from RD 6 to 19, with three instances of the device used to code participants whose RD is 20. One dissimilarity between the figures in Table 5 and those in Table 2 for the VC device is that, especially for O participants, RD is slightly higher for the E device than for the VC device. This indicates that O participants coded with the E device are slightly more discontinuous than A and S_a participants coded with the VC device.

There is no clearly significant patterning of RD indices for the E device according to semantico-syntactic role or character status. For the O role, central characters exhibit a slightly lower RD than do major characters, and major characters exhibit a slightly lower RD than do minor characters. This is yet another example of central characters being more topical than non-central, and major characters being more topical than minor characters. However, the overall spread between the RD indices of the various character statuses is not particularly striking, and the overall RD indices for O and S_o participants of verbal predicates are very similar. I judge that there are not enough examples of S_o participants of non-verbal predicates to draw any conclusions from their RD indices, which are consistently 1.

Table 6 shows the referential distance indices for the post-predicate NP device used to code O and S_o participants:

RD	O				S_o (pred nom)				S_o (verbal)	
	C	MA	MI	T	C	MA	MI	T		Total
1.		1	1	2					(none)	2
2.		2	1	3						3
3.										
4.			1	1						1
5.										
6.		1		1						1
7.		1		1						1
.										
.										
.										
20.		3	7	10			5	5		15
n =		8	10	18			5	5		23
Mn=		9.75	14.7	12.5			20	20		13.83

Table 6: Referential Distance, Post-Predicate NP Device
O and S_o Categories

Table 6 shows that the simple post-predicate NP device is not used to code central characters at all. However, the generalization that major characters exhibit lower RD (9.75) than do minor characters (14.7) still holds. Also, subjects of non-verbal predications (S_o participants) are only used to code participants with the maximum RD (20), though the number of examples of subjects of non-verbal predications coded with this device is so small that any generalizations made with respect to their indices must remain tentative.

Table 7 indicates RD figures for the enclitic plus NP device used to code O and S_o participants:

RD	O				S_o (pred nom)				S_o (verbal)				
	C	MA	MI	T	C	MA	MI	T	C	MA	MI	T	Total
1.	1	5	2	8		4		4					12
2.	2	2	1	5		1		1					6
3.	1	1		2						1		1	3
4.													
5.		1		1									1
6.		1	1	2									2
7.		2		2									2
8.		1		1									1
9.													
10.													
11.	1			1									1
.													
.													
.													
20.	2	3	15	20		1		1					21
n=	7	16	19	42		6		6		1		1	49
Mn=	8.43	6.56	16.32	11.29		4.33		4.33		3		3	10.26

Table 7: Referential Distance, Enclitic Plus NP device
0 and S_o categories

In Table 7 again we see that the only substantial numbers of instances of the E+NP device occur in the O columns (n = 42). Here we see that central and major characters are similar in RD, while minor characters are substantially higher. However, for the first time the generality that central characters are more topical than major characters is violated. In Table 7 we see that central characters coded with the E+NP device are actually less topical in terms of RD (RD = 8.43) than are major characters (RD = 6.56).

3.5.3 Oblique participants. In this section the tables indicating referential distance and figures for oblique participants are presented. The only devices ever used to code oblique participants in my corpus are HC, NP, and HC+NP.

Table 8 presents the RD figures for the HC device:

RD	C	MA	MI	Total
1.	27	41	10	78
2.	5	10		15
3.	2	5	1	8
4.		2		2
5.		1		1
6.				
7.	1			1
8.				
9.		1		1
.				
.				
17.		1		1
n =	35	61	11	107
Mn=	1.43	1.88	1.18	1.66

Table 8: Referential Distance, Head Coding Device
Oblique Categories

From Table 8 we observe that the RD indices for the HC device are roughly comparable to those of the VC (Table 2) and E (Table 5) devices. This is to say that HC is a short-range device in that there is a nice clustering of instances of this device near the lower end of the scale. However, these figures differ from those for VC and E in that there are no instances of the HC device used to code a participant with the maximun RD of 20. In other words, there is no secondary clustering towards the upper end of the scale as there is for VC (see Table 2) and E (Table 5).

Table 9 presents the RD figures for the (pre-head) NP device:

RD	C	MA	MI	Total
1.		3		3
2.		3	5	8
3.		2		2
4.			2	2
5.		1		1
6.		3		3
7.			1	1
8.		2		2
9.		1		1
10.			1	1
11.				
12.			2	2
.				
.				
20.		5	51	56
n =		20	62	82
Mn=		8.2	17.39	15.15

Table 9: Referential Distance, NP Device
Oblique Categories

As with all other roles, the simple NP device used to code oblique participants is reserved for non-central characters.

Table 10 presents the RD figures for the HC+NP device:

RD	C	MA	MI	Total
1.	6	2	1	9
2.	1	2	1	4
3.	3	2	1	6
4.	1			1
5.				
6.		1		1
7.	1	1		2
8.		1	1	2
.				
.				
13.		2		2
.				
.				
17.		1		1
18.		1		1
19.				
20.	1	6	8	15
n =	13	19	12	44
Mn=	3.69	11.26	14.5	9.91

Table 10: Referential Distance, Head Coding Plus NP Device
Oblique categories

In Table 10 we observe that for obliques, central characters are much more likely to be coded with a head coding prefix in combination with an NP than with an NP alone. This fact, along with similar observations made for the other semantico-syntactic roles, allows us to make the more general observation that for all semantico-syntactic roles a simple NP is primarily used to code non-central characters. Central characters are much more likely to be coded with a VC, E, or HC device in addition to the NP.

3.5.4 Summary of mean values by coding device. Table 11 summarizes the aggregate RD figures from all of the preceding tables. In the first column of Table 11 the figures for the three attenuated devices, VC, E, and HC, are found. Since VC is only used for A and S_a participants, the first two boxes in the first column give the appropriate figures for that device. The next two boxes give the figures for the E device, and the last box gives the figure for the HC device. The same is true for column three, where the figures for VC+NP, E+NP, and HC+NP are listed from top to bottom, respectively. At the bottom of these columns the aggregate means for each of these sets of devices are given.

	VC/E/HC	PNP	VC/E/HC+NP	post- NP	All devices
A	1.52	6.2	8.24	----	2.37
S_a	1.7	12.36	7.59	----	3.3
O	2.38	----	11.17	12.33	6.53
S_o	1.53	----	3.75	16.4	5.03
OBL	1.67	15.15	9.91	----	7.97
All roles	1.75	14.37	8.83	13.22	

Table 11: Referential Distance, Summary

3.6 Discussion

Of the three major A and S_a coding devices investigated in this study, I conclude that VC is a 'short-range'[11] (or 'weak') device (mean RD = 1.68). This means that VC is used to code participants that have very recently been mentioned in the discourse, and which therefore are highly topical. NP and VC+NP, on the other hand, are 'long-range' (or 'strong') devices (mean RD = 11.07 and 7.95, respectively). Of the three major O and S_o coding devices, E is a short-range device (mean RD = 2.18) while Post- NP and E+NP are long-range devices (mean RD = 13.22 and 9.98, respectively). Of the three major oblique coding devices, HC is a short-range device (mean RD = 1.67), while NP and HC+NP are long-range devices (mean RD = 15.15 and 9.91, respectively). These facts are consistent with the iconicity principle of topic continuity stated in Sect. 3.1 in that the VC, E, and HC devices are the 'smallest' (i.e., they are the most attenuated devices both in phonological size and in semantic features that they represent) of the major coding devices, and therefore are predicted by the iconicity principle to code the most continuous, least surprising topics. Full noun phrases, on the other hand, should be used to code less continuous topics. That is, when a topic is introduced for the first time, or reintroduced after a long absence from the discourse stage, a semantically highly specified and phonologically large coding device such as a full NP will be needed to code that topic. A small device is likely to be insufficient to distinguish the topic from among all other potential topics available to the hearer.

Some problems to be considered with respect to these figures are 1) under what conditions the short-range devices can be used to code more distant participants, and 2) under what conditions the long-range devices can be used to code recently mentioned participants. These problems will be dealt with in the following section.

4 Exceptions to the general pattern

In this section I will examine specific instances of coding devices that seem to violate the general tendencies observed in Tables 2 through 11. In Sect. 4.1 I look at the thirteen examples of short-range devices used to code participants with RD greater than 8, and in Sect. 4.2 I look at the 113 examples of long-range devices used to code participants with RD less than 4. Many of the observations made in this section are rather speculative, and none is as yet perfectly explicit and predictive. What I have done is categorize the various examples, and suggested possible explanations for their patterning. It may be the case that my explanations can be verified by some psycholinguistic experimental procedure yet to be devised. However, for the purposes of this paper, they will remain simple post hoc observations.

4.1 Short-range devices used to code distant participants

Tables 2, 5, and 8 indicate that there are thirteen occurrences of the short-range devices VC, E and HC used to code participants with an RD greater than 8. Of these thirteen, four are central characters, eight are major characters and one is a minor character. Thus there is a ratio of twelve mentions of central or major characters to one mention of a minor character. In the corpus at large there is a total of 818 mentions of major and central characters to 210 mentions of minor characters, resulting in a ratio of 3.9 to 1. (X^2 with Yates´ correction: 2.82.) Thus status as central or major character is a significant factor in allowing a participant to be coded with a weaker coding device than would otherwise be expected. In the following section I will show that the one instance of a minor character coded with a weaker coding device than expected has an obvious explanation in terms of the context of the utterance. The other anomalous uses of short-range devices are best understood in terms of the hierarchical thematic structure of the text. These factors will be discussed in Sect. 5.1.2.

4.1.1 Contextual inference. TC297 is the only instance of a minor character with RD > 8 coded with a short-range device. It occurs in an episode where the twins are spying on their grandfather to see where he gets water:

(43) Sa-mutá-nuvee-tée sa-dííya-**ra**, ´jpuu, soon.´
3SG-open-ARR2-INTS 3SG-sight-INAN ´gush!´ ´pour!´
´On arrival he opens it in his sight. "Gush! Pour"´ (TC297)

The ´it´ of this sentence refers to some kind of ´spigot´ (as Powlison (1969:115) terms it, though we don´t really know its identity) that Grandfather opens to get his water. This spigot is never mentioned again in the story, and is therefore a very minor prop. In this sentence, the simple enclitic device is sufficient to code the spigot since the verb **muta** ´open´ combined with sound words that can only be used for rushing water make it clear that the thing being opened must be the source of the water. It doesn´t really matter whether the teller of the story really imagined a spigot (which is an item alien to the Yagua culture) or if it is some kind of tree branch or plant or whatever. The actual identity of the item is not important. Rather, it is its function as the source of the water that makes it relevant at this point in the story, and for that purpose the verb and the sound words make the reference as clear as necessary.

4.1.2 Levels of topicality. If the above-mentioned use of a weak coding device is explained in terms of contextual inference, then there are twelve instances remaining to be explained, all of which are either central or major participants. In this section we will see that the fact that all of these examples are either central or major participants is due to the hierarchical nature of topicality. The central characters are topical throughout the discourse, and therefore are highly topicworthy at any given point in the discourse. Major characters are highly topicworthy throughout particular sections of the discourse. Minor characters, on the other hand, come and go and are, in general, unlikely topics at any point. This inherent topicworthiness of central and major characters is enough in the cases cited in this section to overcome a high distance index.

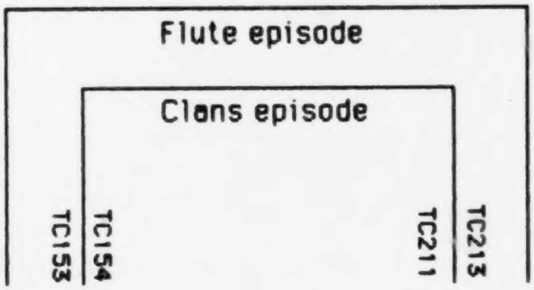
Of course this is not to say that hierarchical structure is more important than distance in determining whether a strong or a weak coding device need be used--in the vast majority of cases distance is clearly the crucial factor. It is only in those few cases where the distance index is overridden that levels of topicality prove relevant. Looking at the twelve remaining instances of central or major characters where referential distance does not explain the use of a short-range coding device, there are particular cases where the hierarchical structure of the text is obviously the relevant conditioning factor. I will illustrate four of these cases here.

In TC213 (see appendix) the following sentence occurs with the simple enclitic device used to refer to the object, the spirit father's magic flute, even though the flute had not been mentioned for sixty clauses:

(44) Santya jɨ́ɨta váriy riiní**rya**.
sa-ntya jɨ́ɨta váriy riy-jiniy-rà
3SG-test JIITA then 3PL-in:presence-INAN
'Then he tested it in their presence.' (TC213)

The flute is then mentioned three more times in succession, but never with an NP. The last previous mention of the flute is in TC153, where the twins test it on their grandmother. Between these two instances of flute testing, however, there is an entire episode in which the twins go about creating all the Yagua clans. Thus the hierarchical structure of the text at this point can be diagrammed as one episode **embedded** in another:

(45)



In TC153 there is a problem to be solved, namely, how to avenge the death of the twins' father. TC154 to TC211 is how the twins go about solving the problem. In TC213 (TC212 is a transition marker) the problem is solved, and the action can continue where it left off in

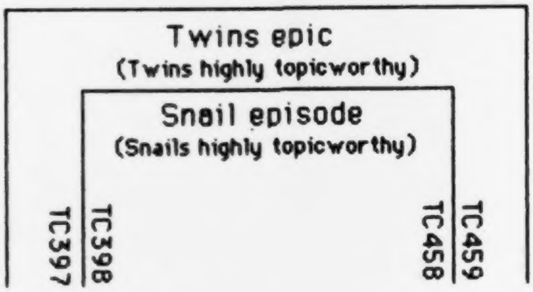
TC153. The flute resumes as topic because its topicality spans the creation of the clans episode.

Another example of topicality spanning an embedded episode occurs in TC459. Here the twins, though the central characters of the entire narrative, have not been on stage for sixty-two clauses, and have not been mentioned at all for sixty-five clauses. And yet in TC459 they are coded with simple verb coding:

(46) Jáschiy jɨ́ɨ́tantíy naaniiniy ránaachǫ pų́ų́riy.
jásiy-siy jɨ́ɨ́ta-ntíy naada-jɨniy rá-naachǫ pų́ų́riy
there-from JIITA-REP 3DL-come INAN-after pifayo
'From there they (two) came after pifayo.' (TC459)

It is significant that the span of text since the previous mention of the twins is clearly definable as an episode. In it the major characters are a water snail and a land snail, and the creation of the Amazon River is explained. The twins are last mentioned in the episode in which they make water available, then comes the snail episode, and then TC459 introduces another episode on how the twins make pifayo available. Clearly the twins, being the highest level topics of the entire discourse, are topicworthy enough at this point to be coded with the simple VC device.

(47)



Ten of the twelve instances of short-range devices coding high RD topics are explainable in terms of the above notion of 'levels' of topicality. Two others are not obviously explainable in this way, though I will contend that they in fact do constitute examples of high-level topics being topicworthy even though their RD is quite high. Each of these happens to be where one of the two twins is introduced into the discourse:

(48) Naanutuvąąchu jɨ́ɨ́ta sųųnááy ruudiimú.
naada-tuvąąchu jɨ́ɨ́ta sa-jųnááy ruudii-mú
3DL-hear JIITA 3SG-cry trash:heap-LOC
'She hears him crying in the trash heap.' (TC22)

(49) Naanutuvąąchuntíy sųųnaantíy.
naada-tuvąąchu-ntíy sa-jųnááy-ntíy
3DL-hear-REP 3SG-cry-REP
'She again hears him crying.' (TC30)

In TC22 Elder Brother is introduced, and in TC30 Placenta is introduced.

We might speculate that even though the twins have not been mentioned prior to this point in the text, they are still highly topicworthy because presumably the hearers know the story well, and understand that the Twins Epic is being recounted, even though the twins themselves have not yet been mentioned. Thus we can consider these clauses to be further examples of higher level topics that are topicworthy at any point in the text. At this point, however, the level of topicality rises to the cultural and social setting of the story itself, with all the activities of the group that had taken place since the last telling of the story constituting an ´embedded episode´ in the ongoing collective awareness of the Twins Epic.

4.2 Long-range devices used to code recently mentioned participants

In this section I will examine specific instances of long-range coding devices (all those involving NPs) used to code recently mentioned participants. The long-range devices distribute much more evenly within the possible range of RD variation than do the short-range devices, i.e., there is a large number (113) of NPs used to code participants with RD less than 4 (see Tables 3, 4, 6, 7, 9, and 10), whereas the number of short-range devices used to code long-distance participants (RD > 8) is relatively small (thirteen to be exact). In the following subsections I will outline seven conditions under which long-range devices are used to code recently mentioned participants, and will give specific examples of each one. In Sect. 4.2.8 I will also present examples that do not seem to fit nicely into any of the other seven categories. The examples presented in these sections were chosen primarily for their brevity. All but one involve long-range devices used to code participants whose RD is 1, allowing most passages cited to be held down to two clauses in length. Hence the examples presented do not exactly constitute a random sampling. However, they do, I feel, adequately represent each of the categories posited.

4.2.1 Ambiguity. Of course a very likely reason for using a strong coding device where referential distance is low is where there are other semantically compatible referents ´on stage´ that might compete for interpretation as the referent of that coding device. Such situations are termed situations of ´interference´ by Clancy (1980) and ´ambiguity´ by Givón (1983a). In the latter work an index of ambiguity is calculated by noting whether or not there are competing referents on stage—if there are no such competing referents, the index is 1; if there are competing referents, the index is 2. Thus a scale of ambiguity ranging from 1 to 2 is defined. Of the 113 examples of strong (long-range) coding devices used to code participants with RD less than 4, fifty-seven were used where the ambiguity was clearly high (i.e., A = 2).[12] The following excerpt illustrates this phenomenon in the First Squirrel text:

(50) a. Naada-rą́ą́y jį́į́ta jás-chiy.
3DL-jump JIITA there-from
´They jump from there.´ (FSQ9)

b. Sa-rą́ą́y jɨ́ɨ́ta múcatyu munátyɨ-ɨ sa-jɨɨsíy.
3SG-jump JIITA squirrel first-NOM 3SG-before
´The first squirrel jumps before him.´ (FSQ10)

In (50a) the squirrel and the deer are collectively referred to with the dual VC form **naadá**, and so in (50b) the RD for both is 1. In (50b) the squirrel is singled out from the deer by the use of a full NP in addition to a VC prefix. Without the NP 50b would be ambiguous as to who jumped first, and the whole point of the sentence would be lost.

In the Twins Cycle, the following passage occurs:

(51) a. Sa-jɨ́ɨ́ jɨ́ɨ́ta sų-ųsíy.
3SG-fly JIITA 3SG-after
´He (Placenta) flies after him (Grandfather).´ (TC296)

b. Sa-múta-nuvee-téé sa-dííya-ra, ´Jpuu, soon.´
3SG-open-ARR2-INTS 3SG-sight-INAN gush! pour!
´He (G) opens it in his (P) sight, "gush! pour!"´ (TC297)

c. Sa-ráni jɨ́ɨ́ta naada-ją́ą́pa rá-áriy.
3SG-stand JIITA 3DL-grandfather INAN-under
´Their grandfather stands under it.´ (TC298)

In (51a) (TC296) the identities of the two participants are clear from the context—Grandfather has just left to bathe, and Placenta has transformed himself into a hummingbird in order to follow Grandfather and find out where he gets water. Again in (51b) the identities of the participants are clear from the context— the one who has gone to bathe is the only one likely to engage in an act involving gushing and pouring, while the other, the spy, looks on. In (51c), however, the context does not help us quite so much to identify the single participant. Yes, we could say that the bather is the only one likely to stand under the flow of water. However, the coding device used to refer to the water here is not explicit. It would not be clear that the thing being stood under is the water flow if the clause did not make explicit reference to the bather. Furthermore, in (51a) and (51b), both Placenta and Grandfather are mentioned in such a way, both pragmatically and syntactically, that precludes coreferentiality between the two mentions within either of these clauses. Thus the hearer knows there are two participants involved in each of these clauses, and must identify one as Placenta and the other as Grandfather. The semantics of these multi-participant clauses is explicit enough to render this task fairly simple, as outlined above. In (51c), however, there is only one participant mentioned, and the hearer must determine whether that one is Placenta or Grandfather. The semantics of this single-participant predicate does help the hearer as much as do the other, more explicit, predicates in (51a) and (51b). Thus the use of a stronger coding device to help the hearer with this identification is justified.

4.2.2 Elaboration. Eleven of the 113 examples of strong coding devices used to code recently mentioned participants are what I will call **elaborations.** Elaboration is where a participant is mentioned in one clause and then further specified, either by noun-phrase modifiers or by a predicate that attributes additional qualities to that participant. One example of this phenomenon is found in KT20:

(52) a. Naada-supátay jɨ́ɨ́ta.
3DL-come:out JIITA
´They come out.´ (KT19)

b. Naada-supáta-myáá jɨ́ɨ́ta-váriy dá-nu-jų́ų́y vánu-jų́ų́y
3DL-come:out-PERF JIITA-then 2-CF:ANIM-2 male-dual
´Two males have then come out.´ (KT20)

In (52b) it is fairly obvious that a full NP is used to code the subject because the subject is being quantified and specified as to gender. This consideration overides the fact that the two males had been mentioned in the previous clause.

A similar example occurs in another text, the Hunter´s Narrative:

(53) a. Sų-ųnúú-ntyíy rá-tiryǫǫ sų-ųnóó.
3SG-see-REP INAN-lie 3SG-head
´He also sees his head lying there.´

b. Jánariy junoo tiryǫ́ǫ́.
deer head lie
´A deer´s head lies there.´

Again it is clear that the head in (53b) is being specified as a **deer´s** head as opposed to any other possible head. For this purpose a full NP is required even though the head had been mentioned in the previous clause.

4.2.3 Discourse promotion. Twelve of the 113 instances of strong coding devices used to code recently mentioned participants are instances of what DuBois (1985 UCLA class lectures) calls **discourse promotion.** This is the phenomenon whereby a participant is first mentioned as an oblique, or possessor, using an explicit coding device such as a full NP. Then in the immediately following clause this participant is again coded with a strong coding device, but ´promoted´ to a more central semantico-syntactic role, i.e., A, O or S. For example:

(54) a. Naan-dííy rí-íva jánariy múdii dárajų́y.
3DL-see INAN-DAT deer jawbone two
´They see two deer jawbones.´

b. Rá-raníy jánariy múdii naana-ajííjų.
INAN-stand deer jawbone 3DL-in:front
´The deer jawbones stood in front of them.´

In (54a) the jawbones are first mentioned as a full NP in the dative case (the verb **dííy** meaning ´see´ requires that the entity seen appear in the dative case). Then in (54b) the jawbones are repeated as a full NP but this time ´promoted´ to the S role.

A similar example occurs in TC70:

(55) a. Si-itǫ́ǫ́-ta jɨ̨ɨ̨́ naadi-imú jí-tyę̨ę̨yada-ra.
3SG-arrive-TRNS JIITA 3DL-LOC 4-grandma-INAN
'He takes it to his own grandmother.' (TC69)

b. Naada-díryey jɨ̨ɨ̨́ sɨ̨-ɨ̨tyę́ę́yadá-yu.[13]
3DL-welcome JIITA 3SG-grandmother-REFL
'His own grandmother welcomes him.' (TC70)

Here the grandmother is mentioned with a full NP in (55a) in an oblique role, and again in (55b) as an S participant.

The phenomenon of discourse promotion illustrates that not all mentions are created equal in terms of activation of participants in memory (Chafe 1985). In particular, it seems that mentions of participants in non-core roles such as obliques do not necessarily suffice to activate a participant to the point where it can subsequently be coded with less explicit devices. This fact suggests a question for further research into coding choice in discourse: how do the various semantico-syntactic roles compare in terms of the degree to which they activate participants in memory? To answer this question we would have to introduce only a minor complication to our methodology for calculating RD indices--in addition to counting clauses back to the last mention of a given participant, we would also record the role of that last mention. Then we could correlate RD with role of last mention. The following is a possible working hypothesis concerning this correlation:

> Let MRD1 be the mean of all RDs between coding device X and the last mention of the participant coded by X where that last mention is in an oblique role.
>
> Let MRD2 be the mean of all RDs between coding device X and the last mention of the participant coded by X where that last mention is in a core role.
>
> MRD2 will be significantly greater than MRD1.

If this hypothesis can be proven, then we would have an empirical basis for the intuition that core roles activate participants more strongly than do oblique roles. The same kind of hypothesis could be devised for any pairing of roles, thus leading to a weighting of the various semantico-syntactic roles in terms of how strongly they activate participants in memory. This would certainly be an interesting and fertile direction for future research.

4.2.4 Dative objects. In both the Hunter's Narrative and in the Twins Cycle there is a pair of central participants that interact throughout the text. When one member of the pair speaks to the other, there is a marked tendency for the addressee to be coded with a full noun phrase, even though that participant may have been very recently mentioned, and even though the noun phrase in question apparently does nothing to disambiguate between the two members of the pair. For example:

(56) a. Jutǫǫ-ją́-ą́siy   si-imú   jí-tyeerɨ-nti-ñíí.
        arrive-o´land-PROX:1 3SG-LOC 4-brother-REP-3SG
        ´He arrives there to where his brother is again.´ (TC304)

     b. Sa-tų́ųchu-nuvee jɨ̨́ɨ̨   sí-íva jí-tyeerɨ.
        3SG-speak-ARR2 JIITA 3SG-DAT 4-brother
        ´He speaks on arrival to his own brother.´ (TC305)

In (56a) one participant is coded explicitly with the NP meaning ´his own brother.´ In (56b) the same NP is used as a dative complement of the verb **tų́ųchu** ´to speak.´ Ambiguity is, technically, high at the point of (56b), since both brothers are on stage. However, the NP in (56b) does nothing to relieve this ambiguity. Since both participants are brothers to each other, the NP could refer to either one. In fact we know from the context that the brother that arrives in (56a) is the one that speaks in (56b), but we could just as easily have come to this conclusion were the brother spoken to in (56b) coded with a less explicit device.

Similar examples occur in the Hunter´s Narrative:

(57) a. Sɨ-ɨmyí-rya ríchanú   naanu-moo-mú.
        3SG-eat-INAN shiringara 3DL-face-LOC
        ´He eats shiringara fruit in front of them.´

     b. Sų-ųta-chí-íva yí-sąą vichɨ̨-ɨ̨́. . .
        3SG-say-3SG-DAT 4-COM be-NOM:ANIM
        ´He says to his own companion . . .´

In (57b) the noun phrase meaning ´his companion´ (literally: ´his own being-with one´, i.e., ´the one who is being with him´, or something like that) does nothing to disambiguate the reference of the addressee. Both hunters are companions to each other, and in this case it could be either one that is speaking. The two hunters are simply not differentiated at this point in the story.

There are eleven examples of this use of full NP for recently mentioned participants in the corpus. The only explanation I have for this phenomenon at present is purely speculative. Since in every case there is a ´semantically appropriate´ referent in the immediate context, perhaps the speaker feels constrained to use a device normally used in situations of high ambiguity, even though in these particular cases a full NP does not accomplish the task of disambiguation. Under this analysis, these examples would be additional examples of the use of full NP in situations of high ambiguity (Sect. 4.2.1).

4.2.5 Thematic structure. Clancy (1980) shows how coding choices in English and Japanese narratives are at least partially influenced by the thematic structure of the text. In particular, noun phrases tend to be used at thematic junctures, even though RD may be low. The particular thematic boundaries that Clancy considers are ´world shifts,´ i.e., shifts between the ´real world´ where the narrator and an interviewer are the participants, and the ´story world.´ Clancy also considers ´episode boundaries.´ Fox (1984), working primarily in the framework of ´story grammar´ (Rumelhart 1975), makes similar observations for

English. In a similar vein, Derbyshire (to appear) shows that free noun phrases in addition to verb coding devices are used at points of thematic discontinuity in Hixcariana, a language of the Brazilian Amazon region.

The general observation that thematic junctures are sometimes accompanied by stronger than otherwise necessary coding devices is also relevant for Yagua, though the nature of the units and boundaries that are especially relevant to Yagua is still a matter for further investigation. In the following excerpt Squirrel is referred to with a modified NP, ´the one who makes him jump´ even though he is mentioned in each of the eight previous clauses, and is therefore a highly topicworthy participant:

(58) a. Squirrel: ´Yi-núúy rą-ą rą́ą́-kyu.´
2SG-see 1sg-IRR jump-potential
´"You see, I can jump!"´ (FSQ11)

b. Sa-niy suvų́-tyéé jiñu munátya sų-ųmutąąsá
3SG-MALF fear-INTS this ancestor 3SG-behind

játiy sa-rąą-ñíí.
REL 3SG-jump-3SG

This ancestor (the deer) is really afraid behind
the one that makes him jump. (FSQ12, 13)

c. Deer: ´Rą-ą jų́ų́-charatá jiyu-dáy koodí-vyiimú.
1SG-IRR fall-might here-DAY snake-inside
´"I might fall here inside a snake."´ (FSQ14)

However, when we look at the story structure of the text, we notice that FSQ12 (58b) occurs at a fairly major thematic boundary. Up to FSQ11 (58a) the theme is, ´the squirrel tries to trick the deer into crossing the stream on the back of the boa.´ The theme beginning in FSQ12, however, is, ´the deer debates within himself.´ There is an obvious shift from the external actions of the squirrel to the internal state of the deer. In terms of Rumelhart (1975), this shift corresponds to an ´event´ boundary where the sequence beginning in FSQ12 (and continuing for several clauses) is a ´reaction´ to the sequence ending in FSQ11. Of the 113 examples of long-range coding devices used to code recently mentioned participants, four are explainable in terms of some notion of thematic boundary, though the exact character of those boundaries is still to be defined. Not all are as clearly related to story structure as (58) is.

4.2.6 Crucial inanimate participants. In the Twins Cycle there are three examples of inanimate participants introduced into the discourse with multiple full noun phrase mentions in succession. In each of these cases, the inanimate participant involved is one which is particularly salient in the subsequent episode. Each of the three instances is at the beginning of a major episode of the Twins Cycle: 1) how the twins obtain water, 2) how the twins obtain pifayo, and 3) how the twins obtain corn. As might be expected, water, pifayo, and corn are the salient inanimate participants involved:

(59) a. Sa-tada-chu-muu-myáá jɨ̨ɨ̨ta Rísu
3SG-recede-cause-COMPL-PERF JIITA God

naana-jisi-ntí-rya ją́ą́-ntíy.
3DL-from-REP-INAN water-REP

´God had also caused the water to recede from them.´
(TC272)

b. Néé ją́ą́-dana-ñuudáy.
NEG water-not:exist-any:more
´There is no more water.´ (TC273)

c. Múú-chi-ñumaa naada-ją́ą́pa-mu-síy
there-from-now 3DL-grandfather-LOC-from

jí-ryi-yąą-ra ją́ą́.
4-get-DIST-INAN water

´Now they repeatedly get water from their grandfather.´
(TC274)

According to the definition of the various character statuses (Sect. 3.3), water is a minor character, since it is never mentioned more than four times in any stretch of twenty clauses. However, it is clearly a major participant in this episode, since the whole episode is about how the twins obtain water. The use of three full noun phrases in a row to code the water iconically represents the importance of water to the episode. It is interesting that crucial animate participants are not introduced in this way. In fact, the Twins themselves, in the same version of the same story, are first mentioned with simple verb coding (TC22 and TC30, see Sect. 4.1.2). This observation is understandable in terms of the fact that inanimates are less likely to be important participants than are animates. When important animates are introduced, no special coding other than whatever is needed to make their identity clear is necessary. For important inanimates, however, special coding is necessary in order to signal that something unusual is going on. Whereas mention with a single NP may suffice to activate an animate participant for future deployment, inanimates, by virtue of the fact that they don´t characteristically persist as important participants, need the reinforcement of several full NP mentions in order to be sufficiently activated in the hearer´s memory.

The following sentences introduce the episode in which the Twins obtain pifayo from their stingy Grandfather:

(60) a. Jás-chiy jɨ̨ɨ̨ta-ntíy naani-iníy rá-naachǫ́ pų́ų́riy.
there-from JIITA-REP 3DL-come INAN-towards pifayo.
´From there they come looking for pifayo.´ (TC459)

b. Pų́ų́ri-vya-numaa jɨ̨ɨ̨ta naada-yasanta-ntí-ñíí.
pifayo-DAT-now JIITA 3DL-pester-REP-3SG
´They now pester him for pifayo.´ (TC460)

Here pifayo is mentioned twice in a row with a full NP. This is not technically speaking an instance of discourse promotion, since both mentions of pifayo are in oblique roles. Also it is not elaboration, since there is no modification of the reference of pifayo in the second sentence. Like water in the previous set of examples, pifayo in (60a) and (60b) goes on to be a significant element in the subsequent development of the episode introduced in (60a). Of the 113 examples of long-range devices used to code recently mentioned participants, five are explained in terms of this notion of 'introduction of crucial inanimate participant.'

4.2.7 After mention in quotes. There are four examples in my corpus of full noun phrases being used to code participants that had been mentioned in a quote in the immediately preceding clause. Because of my method of counting mentions of participants within quotes as full mentions, the RD for these instances is 1. For example:

(61) a. 'Ra-chikidi-naachǫ́ǫ́ vuryą-ą junúú-yąą-téé-kii.
1SG-intestines-towards 1PLINC-IRR look-DIST-INTS-must
'"We must look all around for my intestines!"' (TC517)

b. Díy ri-jyetyą́-ą́siy ri-inúú-rya.'
there 3PL-throw-PROX1 3SG-see-INAN
'"There I saw them throw them."' (TC518)

c. Sa-ryiy jɨ́ɨ́ta jí-chikidi.
3SG-get JIITA 4-intestines
'He gets his own intestines.' (TC519)

In (61c) the intestines are referred to with a full NP even though they were mentioned in the preceding two clauses, and there are no other inanimate participants cluttering the discourse stage at this point. My conclusion is that perhaps my method of counting mentions within quotes as full mentions needs to be revised. It seems from examples such as this that mentions of participants within quotes, like mentions of participants in oblique phrases, are not as salient as mentions of participants in core semantico-syntactic roles in straight non-quote clauses. This hypothesis could be tested in a manner similar to that outlined in Sect. 4.2.3 for oblique mentions.

4.2.8 Residue. Finally, there are nine examples of long range devices used to code recently mentioned participants that do not seem to fall into any of the above categories, and which don't have any obvious idiosyncratic motivation. A couple of these will be presented here:

(62) a. Sa-deenu-dee-rá yí-náá-tyéniy
3SG-child-DIM-NOM:NEUT 4-cry-cause

váácha jii-tąąsá.
huapo:monkey branch-middle

'The huapo monkey makes his own child cry in the middle of the branch.'

b. Są-ąta-tyén-níí tápi raatya-déé jíí-va.
3SG-jump-cause-3SG slow careful branch-DAT
´He makes him jump slowly, carefully along the branch.´

In (62b) there is no obvious reason why the branch is coded with a full NP. This example is somewhat akin to discourse promotion (see Sect. 4.2.3) in that the reference to branch in (62a) is in an oblique role. However, there is no ´promotion´ involved since the reference to the branch in (62b) is also in an oblique role. Nevertheless, the explanation of discourse promotion in terms of less salience ascribed to mentions of participants in oblique roles would still potentially be relevant here, i.e., the branch is referred to with a full NP in (62b) because the speaker judged the mention of the branch in (62a) to insufficiently activate that participant for further non-explicit mention.

Another unexplained example of use of a full NP occurs in KT45:

(63) Rá-jųų-yaa-mű́úy naada-sųųtá.
INAN-fall-DIST-COMPL 3DL-shelter
´Their shelter fell to pieces.´ (KT45)

At this point in the story the shelter has an RD of 3, where the last mention of the shelter is in the O role. There are no potentially interfering inanimate participants on stage, and there is no other obvious reason why the speaker did not use a less explicit coding device. Hence this is an ´unexplained´ example. It may be the case that since this is an inanimate participant, and not a very salient one at that, an RD of 3 is sufficient for it to decay from active memory, whereas animate and otherwise more salient participants decay more slowly. This observation suggests further hypotheses which will have to await future research.

5 Conclusion

In this paper I have presented the results of a quantitative study of the use of certain participant coding devices in a body of folkloric texts. The specific measurement employed is that termed ´referential distance´ (Givón 1983a,c,d), though several modifications to the methodology as outlined by Givón are proposed. These modifications relate primarily to the treatment of quoted material. The general findings of this study confirm the iconicity principle of topic continuity as expressed in Givón (1983a), though there are several individual instances that deviate from the general pattern. In Sect. 4.1 and 4.2 I look more closely at these exceptions and class them into two general types: 1) short-range devices used to code relatively distant participants, and 2) long-range devices (i.e., all those involving full NPs) used to code recently mentioned participants. In the first class, most examples are explained in terms of the hierarchical nature of topicality, i.e., participants are highly topicworthy (and therefore codable with attenuated coding devices) throughout the span of text in which they figure prominently, even though they may cease to be mentioned for long stretches within those spans. In the second class, most examples are explained in terms of the index of ambiguity, i.e., the presence of other semantically appropriate participants in the immediate context. In addition to the examples explainable in terms of ambiguity, there are several other subclasses of examples of long-range devices used to code recently mentioned participants. These subclasses are discussed in Sect. 4.2.2 through 4.2.7, and tentative, sometimes speculative explanations are provided. Finally in Sect. 4.2.8 some examples of unexplained uses of long-range devices are presented.

This study shows that the quantitative methodological procedure coming to be known as 'topic continuity' is a useful base upon which to build observations concerning the use of coding devices in discourse. However, much work is yet to be done before a truly predictive theory is available. This study has attempted to refine the methodology, and to point out areas where further research is in order.

Appendix

The Non-Identical Twins Cycle (TC)

From Powlison (1969) and Dorson (1975)

Laureano Mozombite

The following is a retranslation (by Mamerto Macahuachi and T. Payne) of a tale that appears in Yagua in Powlison (1969:176-219), and in English in Dorson (1975:553-6). All translation mistakes in this version are my own responsibility. Like many long tales, this story is a series of episodes, each of which may be told as individual stories. Powlison (1969) provides a detailed description of the various Yagua epic tales, and how the episodes and individual characters intertwine. The overarching theme of this tale is how the world came to be the way it is. Some of the specific topics dealt with in this version are 1) the significance of the distinction between the Yaguas and the non-Yagua Indians, 2) how the Yagua clans were created, 3) how the Amazon river was created, 4) why life must involve difficult labor, 5) the origins of blowguns and other hunting equipment, 6) how water turtles and land turtles came into being, 7) how pifayo (**guilielma gasipaes**, a palm fruit) was obtained by the Yaguas, and 8) how corn was obtained by the Yaguas.

This tale forms part of the data base for the topic continuity study described in this paper. Although it would be preferable to include the full Yagua transcription of this text, due to space limitations I have limited presentation to the English translation. The English is included in order to help the reader understand something of the thematic structure of Yagua discourse and to locate specific examples from this text cited in the body of the paper in their discourse contexts. As mentioned above, this text appears in Yagua in Powlison (1969), though in a different orthography than that which is employed in this paper, and without clause numbering.

1. Yes, Creator created. Creator created long ago.

Scene I: The House

---

2. The adults drink manioc beer,
3. beer they drink.
4. Her pregnant daughter says to the old woman:
5. ´While you weed the manioc patch, we are going to continue drinking.
6. You don´t drink anything with us.´

———— Transition ————

7. She goes.
8. They drink in her absense.

Scene II: The Manioc Patch: Grandmother Worries

---

9. After a while, suddenly they are quiet, silent.
10. She listens and listens.

11. 'What could have happened to them again, what?'
12. They aren't laughing, they aren't drumming anymore.
13. It's almost evening.
14. Finally she goes.

Scene III: The House: Grandmother Discovers Carnage

---

15. She looks on arrival. 'jiii'
16. The house is smoking, the ruins of the house.
17. The savages have burned it.
18. 'Certainly the savages have completely killed them!
19. No wonder they're not drumming, no wonder!
20. Clearly they've all been killed!'
21. From there she is wandering around
22. when she hears crying from the trash pile:
23. 'Cuway, cuway, cuway, cuway!'
24. 'Jiii,' here clearly the savages have thrown my daughter's child!'
25. She goes.

Scene IV: The Trash Pile: Grandmother Discovers the Twins

---

26. 'I will recover him to be my companion.
27. I will raise him to be my companion.'
28. She recovers him.
29. As she is going, she hears that another is crying there also.
30. 'Is someone there?'
31. She returns again.
32. It is his placenta that has been transformed.
33. She recovers him also.
34. She goes, then, under her shelter again.

Scene V: Grandmother's Shelter (in the Manioc Patch): The Twins Grow

---

35. She washes him there on arrival.
36. In two days they sit up.
37. In three days they walk all over the place.
38. They don't delay in growing.
39. In five days they are complete adults.
40. He asks his grandmother:
41. 'How then did my deceased father die, how?
42. And my deceased mother?'
43. 'The savages just killed them.'
44. 'Really?'
45. 'Yes.'

Scene VI: The House: The Twins See Game

---

46. The two go again there in the neighborhood of the ruined house of their father,
47. and they pass by all sides,
48. and they see little toucans, everything: toucans, wild turkeys.

49. They are eating tayra berries.
50. ´What can we use to kill them?´
51. The two return to their grandmother again.

Scene VII: Grandmother´s Shelter: The Twins Learn about Hunting

52. ´With what did my father hunt animals, with what?´
53. ´With just a blowgun.´
54. It´s not here, you know, the blowgun,
55. that which is its tree (that which is the blowgun tree).
56. From there your deceased father got his blowgun.
57. Darts also from the fork (heart) of the inayuga palm he got them.´
58. When it dawned again,

Transition

59. they left for the tree.
60. ´Be careful it traps you!
61. Quickly you must snatch it from inside,
62. if you want to get a blowgun.´
63. They go to the blowgun tree.

Scene VIII: The Blowgun Tree: The Twins Get a Blowgun

64. It is yawning over and over again in front of him, ´Po, po, po!´
65. Right close by he is now, right close by.
66. There grabbing it he yanks, ´Siyon!´
67. There it springs out beside it.
68. So he grabs his blowgun.

Scene IX: Grandmother´s Shelter: The Twins Learn about Blowguns

69. He carries it to where his grandmother is.
70. He greets his grandmother;
71. ´Why have you ruined it again?´
72. ´Why not?´
73. ´For what purpose do you ruin it, for what purpose?´
74. ´So that our offspring will have to suffer (work hard) to make their blowguns.
75. Isn´t it important that they make them with their hands? (rhetorical question meaning ´you know it´s important that they make them with their hands´).

Transition

76. They go again for darts.
77. ´Be careful, the scorpions that protect it bite you!
78. It´s not just one biting thing that protects it, red scorpions and snakes also.

Scene X: The Inayuga Palm: The Twins Get Darts for Their Blowgun

79. He climbs searching to the fork of the inayuga.

80. He finishes off the scorpions, the red scorpions, the snakes,
81. and collects from where they were the darts.

Scene XI: Grandmother's Shelter: The Twins Learn about Darts

---

82. She sees also that he carries a roll of darts.
83. 'Why did you finish off the biting things that protected it also?'
84. 'Why not?'
85. 'For what reason did you finish them off, for what reason?'
86. 'So that they will have to whittle their darts with a knife.
87. They go again for a dartholder,

Scene XII: The Catirina Palm: The Twins Get a Dartholder for Their Darts

---

88. which is in the fork of the catirina palm.
89. There are biting things that protect it also.
90. He finishes them off,
91. the ones that protected it also,
92. and he gets the dartholder also.

Scene XIII: Grandmother's Shelter: The Twins Learn about Dartholders

---

93. His grandmother sees also.
94. 'Why did you finish off the biting things that protected it also?'
95. 'Why not?'
96. 'So that they will have to weave their own dartholders.'
97. Thus it remained there (i.e., like that).

Scene XIV: The House: The Twins Hunt and Grab the Magic Flute

---

98. The two of them go from there again around the neighborhood of the ruins of the house of their deceased father.
99. There they went blowgun-hunting little toucans, everything, wild turkeys.
100. There they spy on the spirit of their mother, and the spirit of their father,
101. those that dance in the middle of the ruins of the house.
102. Another day, the same thing again.
103. Another day, the same thing again.
104. Finally the two think:
105. If only we could snatch the flutes of our deceased parents,
106. with them we could secure vengeance for our departed parents.

107. Finally he says to his brother:
108. ´Let´s go and snatch them!´
109. The two of them get up early again.
110. Today, yes, we are going to snatch them.
111. They hide nearby, there where they circle (all traditional Yagua dancing involves circling).
112. ´Here is where they come circling.´
113. They put cetico leaves over themselves, the other one too.
114. ´Be careful not to let go when you grab it!
115. If it heats up,
116. your hand will heat up immediately also.
117. If it shrinks to a tiny flute,
118. your hand will shrink immediately also.
119. If it enlarges to a huge flute,
120. your hand will enlarge immediately also,´
121. he says to his brother.
122. The two (spirits) descend to earth again.
123. They are dancing on arrival,
124. ´Puju!´ they dance.
125. There he jumps up suddenly against them.
126. ´Rupa!´ His brother has grabbed his deceased father´s flute.
127. ´Tanti!´ The two of them stick together.
128. His brother now screams:
129. ´Now my hand is burning!´
130. ´Equally your hand will heat up immediately also.´
131. ´Now it burns me!
132. Now it burns me!
133. Now I let go of it!´
134. ´Don´t let go of it!´
135. Finally, he yanks his flute from him completely.
136. Only Placenta succeeds in grabbing the flute of his deceased mother.
137. Finally he snatches it away from her,
138. and she ascends jumping, ´puri.´
139. The two have snatched the flute from their deceased mother.
140. The two go again there to where their grandma is.

Scene XV: Grandmother´s Shelter: The Twins Test the Flute, Grandmother Makes Manioc Beer

---

141. They call to their grandmother upon arrival:
142. ´Grandmother! Here, sit down.´
143. She sits down at their command,
144. and they blow into it, ´Vɨɨ, vɨɨ, vɨɨ, vɨɨ, vɨɨ, vɨɨ.´
145. One strong blow into it, ´kiiin.´
146. She falls like dead.
147. Their grandmother revives again.
148. She says,
149. she scolds her two grandsons:
150. ´Jii! What´s happening with you two?
151. Obviously you´ve succeeded in grabbing the flute of

your deceased mother also!´
152. ´Why not?
153. Maybe with it you (speaking to other twin) will secure vengeance for our dead father.´
154. He says to his grandmother:
155. ´Grandmother, prepare manioc beer.´
156. ´Who, then, is going to drink it with you, who?
157. There aren´t any people that I see (i.e., know about) around here, none.´
158. ´Just nearby there are people.´
159. ´(But) the house is not large enough.
160. Where then will the people sit, where?´
161. ´I´ll just make the house larger then!´
162. She makes the manioc beer.
163. ´Jujum,´ she finishes making it.
164. When she finished making it,
165. she says to her grandson:
166. ´Here now is the manioc beer that you requested.´
167. He commands then his elder brother:
168. ´Invite!´
169. ´Where then will I go to invite, where?
170. I have not seen any people around here, none.´
171. ´But just over there there are people.´
172. Finally, he goes.

Scene XVI: The Jungle: Elder Brother Searches for People

---

173. From there he goes circling,
174. without seeing any people,
175. and returns.
176. ´Aha! Many have you invited?´
177. ´I didn´t see anybody to invite.´
178. He greets his brother:
179. ´Where, then, do people lack, where?
180. I then will invite!´
181. Placenta goes.

Scene XVII: The Jungle: Placenta Creates the Clans

---

182. Near the top of a heap he kicks.
183. ´Come and drink at my place!´
184. From there on the top of a heap of Macaw feathers,
185. ´Jun! Come and drink at my place!´
186. From there also on top of a spotted cavy burrow.
187. ´Come and drink at my place!´
188. From there also upon an ant´s nest.
189. ´Come and drink at my place!´
190. From there also against the buttress root of a pachaco tree he kicks in passing also.
191. ´Come and drink at my place!´
192. From there on top of a bat´s nest he kicks in passing also.
193. ´Come and drink at my place!´
194. From there against the trunk of the blowgun tree he strikes in passing.

195. ´Come and drink at my place!´
196. He turns back from there.
197. ´Enough now.´

Scene XVIII: Grandmother´s Shelter: The Clans Arrive and Drink

198. He says upon arrival to his brother again:
199. ´Do you think I have invited (things) which you say are people?´
200. There are no people within a great distance, none.
201. After a long time, a long time, they now arrive in groups.
202. Those of the Squirrel clan begin to arrive,
203. those of the Red Macaw clan after them,
204. those of the Spotted Cavy clan after them,
205. those of the Ant (Isula, a large stinging ant) clan after them,
206. those of the Pachaco tree clan after them,
207. those of the Bat clan after them,
208. those of the Blowgun clan after them.
209. And the house was filled with people.
210. They pulled out the supports of the house to make a large house.
211. They drink all night long, all night long.
212. It dawns.
213. He tries it on them,
214. (to see) if it works.
215. He blows into it in the midst of them.
216. He blows hard into it.
217. They all fall then.
218. So he says:
219. ´It works!´
220. So they get up early (to fight) against the savages.
221. ´Let´s go to the savages!
222. Let´s kill them!´
223. They go.

Scene XIX: The Savages´ House

224. They arrive near the house of the savages.
225. Placenta transforms himself into a small hawk, beautifully speckled.
226. He ascends running along the roof on arrival,
227. because the savages´ roof didn´t reach the ground.
228. He ascends running along the roof on arrival.
229. The savages hear:
230. ´Who then is running up there on the roof?´
231. They come out.
232. They look.
233. ´Who then also?
234. How beautiful!
235. Look!´
236. He calls to his kinsmen.
237. They come running out.

238. ´Jiin, unreal!´
239. They all run out of house, all the savages.
240. One carries a blowgun.
241. ´I´m still going to blowgun him!´
242. ´Don´t blowgun him yet!
243. We should tell the people of the other house too,
244. that they might come right away.´
245. One runs to the other house
246. and tells them also.
247. They ran then also. ´Jiin.´
248. He begins to blow into the flute of his deceased mother, ´Vɨɨ, vɨɨ.´
249. He says to him:
250. ´Don´t blowgun him yet.
251. How is he going to play (the flute)?´
252. ´O.K.´
253. He is blowing into it.
254. He blows. He blows. He blows. He blows.
255. Quickly he now blows.
256. With all his strength he blows into it.
257. ´Yun!´ All the savages fall over,
258. there they are laid out.
259. Not one remains (standing).
260. So the ones that were with him run in passing.
261. There they kill with clubs.
262. ´Juuuun.´ Finished!
263. ´Your request.
264. That´s how they killed my deceased father.´
265. The matter was finished.
266. They turn back after the battle.
267. They drink the leftover manioc beer in the house again.
268. That´s how all the clans remained.
269. So they all stayed outside.
270. So it was he who created the Squirrel clan, the Red Macaw clan,
271. so he created them all.

Scene XX: The Twins Obtain Water

---

272. God has caused the water to subside from them until it is all gone.
273. There isn´t any water anymore.
274. From then on they keep on getting it from their grandfather.
275. Day by day, day by day,
276. ´Tiiy,´ until they are tired of it.
277. ´I´m tired of this!´
278. They ask one who lives there with their grandfather,
279. ´How does he get water?´
280. ´I don´t know.´
281. ´Don´t you know where he bathes?´
282. ´He always goes bathing over there.
283. He bathes at noon.´
284. He says to his placenta again,

285. ´Go see where.
286. This fellow says he bathes at noon.´
287. ´O.K.´

Scene XXI: Grandfather´s Bathing Place

---

288. He goes to the edge of the woods
289. and watches patiently from there.
290. At last he (Grandfather) speaks,
291. ´Ugh! It´s too hot for me!
292. I´m going to bathe first.´
293. The sun is directly overhead.
294. He (Placenta) goes then
295. and changes himself into a little hummingbird
296. and flies after him.
297. He (Grandfather) opens (a spigot) when he gets there in his sight (Placenta´s), ´Gush! Pour!´
298. Their grandfather stands under it.
299. The hummingbird is flying along,
300. ´Tu tu tu.´ He hits at him.
301. ´Hummingbird, Hummingbird! Why are you being a nuisance?´

Scene XXII: Back Home

---

302. There he returns ´Chiy! Chiy! Chiy! Chiy!´
303. He returns.
304. He arrives back at his brother´s,
305. and tells his brother,
306. ´It´s in that whatcha-ma-call-it water tree which is standing,
307. that great big tree standing (there)!´
308. ´Really!´
309. ´Yes.´
310. ´What shall we do?´
311. ´I don´t know
312. unless we cut it down.´
313. ´O.K.!´

Scene XXIII: Grandfather´s Place

---

314. They rise and go early the next morning to their grandfather´s again.
315. They say to their grandfather when they get there,
316. ´Grandpa?´
317. ´What?´
318. ´Uhh, we´re going to cut down this tree which is standing.´
319. ´Go ahead and cut!
320. It isn´t forbidden to cut it down.´
321. They invite (to work) with them woodpeckers, squirrels, agoutis—all of them—
322. woodcreepers, those who make holes, barbets.
323. They invite them all (to work) with them.
324. They cut it.

325. They begin cutting
326. and cut and cut and cut,
327. as far as its center.
328. They cut and cut
329. until the woodpecker is into its heart.
330. 'Ti!' It's getting late.
331. It's late.
332. They give up on it.
333. It is quite thin (when) they leave off.
334. 'Tomorrow we'll fell it!'
335. 'Yes.'
336. It dawns
337. and they go again.
338. It stands there intact again.
339. 'No doubt he put its chips back again!'
340. They cut again. 'Ti!'
341. They cut and cut and cut
342. until it isn't very thick anymore.
343. At last he sends his placenta again,
344. 'Go listen, transforming yourself into the likeness of a little bird,
345. to what Grandfather says.'
346. Their grandfather is sitting in the yard, smoking.
347. He (Placenta) goes.
348. He transforms himself.
349. He listens.
350. He (Grandfather) smokes,
351. he blows it around.
352. He (Grandfather) speaks and
353. he (Placenta) hears,
354. 'Those two children will never fell the whatcha-ma-call-it water tree!
355. They'll never be able to fell it,
356. unless they should make a scorpion bite the tip of my little toe.
357. Then, it would fall.'
358. Then Placenta turns back again.

Scene XXIV: Back Home

358a. He tells his elder (brother),
359. 'This is what Grandpa said,
360. only if we were to get a scorpion to bite the tip of his little toe,
361. then it would fall.''
362. He says to him,
363. 'Transform yourself then!'
364. 'Into what?'
365. 'Transform yourself into a scorpion.'
366. He transforms himself again.

Scene XXV: Grandpa's

---

367. He has gone again.
368. He (Grandfather) smokes, and smokes,

369. and blows it around on (the tree's) vines.
370. There he bites him on the tip of his little toe when he gets there, 'siii.'
371. 'Hey!' It begins to crack immediately,
372. 'Yikes! Ouch! How wise these two kids are!'
373. It stays,
374. it stays just a little bit on the lean now.
375. He says to him again,
376. 'Who on earth is the most painful biter?
377. 'The red scorpion, I suppose.'
378. 'Transform yourself into a red scorpion.'
379. He transforms himself into a red scorpion.
380. He goes again
381. and bites him the same way on arrival again.
382. He was indeed a very painful biter.
383. 'Sii.' He falls over,
384. 'Pųų.' It falls then,
385. 'Yųųųn, pųų.'
386. It falls.
387. His grandson runs at the same time to him,
388. 'What happened to you, Grandpa?'
389. He is not alive anymore.
390. He had died.
391. He blows on him, 'Jųųųų.'
392. He sits up,
393. 'Ha! How are you two so wise?
394. No doubt you've cut it down, too!'
395. 'Of course.
396. What is our posterity supposed to drink?'
397. 'O.K. Let it be so!'

Snail Episode (No Scene Change)

398. A little snail comes running for a leaf,
399. and grabs it for his door plug.
400. He touches (it).
401. In his view, it makes a pretty sound.
402. The first water snail comes running to him.
403. He says then to him,
404. 'You just got that?'
405. He asks him for it.
406. 'Let me see!'
407. He gives it to him then,
408. 'Go ahead and look at it.'
409. He handles it with his hands.
410. 'How very pretty it is!'
411. He gives it back to him,
412. 'Here it is!'
413. He asks for it again.
414. He gives it to him again.
415. He rubs it in his hands.
416. At last he says to him departing,
417. 'I have it now!'
418. He runs away from him to the water.
419. He jumps with it away from him into the water.

420. The owner of the taken object follows him for it.
421. He has jumped with it into the water ahead of him.
422. He jumps in after him.
423. ´Tííy,´ he can´t submerge.
424. He just floats around.
425. The land snail speaks then,
426. ´Why did you impoverish me?!
427. Now the isulillo (ants) will always bite my exposed fleshy parts.´
428. Another comes running, too,
429. ´I´ll be a water tortoise.´
430. He jumps into the water, ´Tápɏɏ.´
431. ´Yuu,´ he floats up.
432. ´Tííy. You can´t be a water tortoise!´
433. Another comes running,
434. ´I´ll be a water tortoise!´
435. In he jumps, ´Tápɏɏ.´
436. ´Piri.´ He sinks,
437. the one who says he will be a water tortoise.
438. He lands on the bottom of the Amazon.
439. ´I´ll be the water tortoise,
440. you be a land tortoise.´
441. That one remains as the water tortoise.
442. ´I´ll be the water tortoise,
443. you be the land tortoise.´
444. ´What will I eat?´
445. ´Well, fungus and tortoise fruit.
446. You´ll eat the tortoise fruit.
447. The berries which ripen red.´
448. ´O.K.´
449. Its owners (the caterpillars that ´own´ the water tree) are now transformed one by one.
450. They paddled away.
451. The white people paddled away, as whites, as blacks, all of them, Cocama Indians.
452. All its owners go transformed.
453. Its chips have all been transformed into fish, which are the umbrella tree chips.
454. All its leaves transform into what they call mojarra fish, a long kind of mojarra fish.
456. They all transform.
457. Gamitana fish, arapaima, all its leaves are transformed into fish.
458. It (the tree) became the long Amazon River then.

Scene XXVI: The Twins Obtain Pifayo

---

459. From there they (two) come for pifayo.
460. Now for pifayo they pester him again.
461. ´Give me a pifayo seed, Gramps!´
462. He gives him a whole stalk.
463. They cut them all in half.
464. ´Tííy.´ None have seeds.
465. ´None of them have any seeds.´
466. He gives them another stalk.

467. ´Tííy.´ Neither does this one.
468. ´This one doesn´t either.´
469. Finally, he thinks again.
470. ´Let´s steal it.
471. Just over there is one that has seeds.
472. He always gives us ones that don´t have seeds.´
473. He goes.
474. Transform yourself.
475. Transform yourself into a parakeet.´
476. ´O.K.´
477. He calls the parakeets to himself.
478. Many he calls.
479. He goes from there to his grown-over garden,
479a. there where his pifayo palm grove is.
480. ´This is really thick.´
481. The parakeets descend on it, ´Yʉʉʉʉ.´
482. There they are destined to be blowgunned by him.
483. They pile into the pifayo.
484. So there he shoots (blows).
485. So then he shoots.
486. So then he shoots.
487. There now Placenta penetrates it.
488. So then he shoots.
489. There at last he finds its seed.
490. When he comes out from there,
491. there he immediately shoots him, ´Kɨɨɨ.´
492. ´Ke, ke, tíye,´ he falls.
493. ´Yʉʉʉ,´ so then they scatter all over from him.
494. He gathers his kill.
495. ´Jʉʉʉ,´ its a big pile.
496. He takes them.
497. They defeather them when he arrives.
498. They clean them.
499. They cook them in pifayo peel water.
500. Then his brother arrives.
501. They take the parakeets from the fire now.
502. His grandfather gives him to him.
503. ´Eat this one at your house.
504. They are no good pifayo wreckers.
505. He was wrecking pifayo.
506. It´s his BROTHER he gives to him.
507. He says to him,
508. ´Quick! Hurry!´
509. He whispers to him.
510. ´Lick my eye.
511. Pifayo water has already entered my eye.
512. Let´s go!´
513. He says to his grandfather,
514. ´I´m going now, Gramps.´
515. ´O.K.´
516. There at the forest edge he says to his brother,
517. ´Let´s look around for my intestines.
518. Yonder I saw them throw them.´
519. He grabs his intestines.
520. He inserts them inside himself.

521. Then he is transformed into a person again.
522. He gives the pifayo seed to him.
523. 'This is the pifayo seed you requested.
524. It is just a little bit cooked.
525. Try planting it.'
526. He plants it.
527. It dawns.
528. It's already this size of tree at dawn.
529. After three sleeps, it already gives fruit.

Scene XVII: The Twins Obtain Corn

---

530. There corn is now lacking.
531. Not just recently they keep asking their grandfather for corn also.
532. He doesn't want to give them any.
533. Finally they say to the grasshopper,
534. 'Can you steal one seed for me?'
535. 'Sure.'
536. He roasts coca leaves.
537. Grasshopper roasts them.
538. Their grandfather is just sitting outside.
539. Finally he stands up.
540. The coca leaves now stir themselves around in his absence.
541. 'Yǫ, yǫ, yǫ.' As he shells (corn),
542. it spills all over, 'yųųųų.'
543. 'Grasshopper, Grasshopper, you are stealing again!'
544. 'Not at all. I'm just here roasting coca leaves.'
545. He (GF) goes after it (the spilled corn),
546. and puts it all back where it was.
547. He puts it back where it was.
548. 'Pųų,' another grain also, another grain also. (grain by grain).
549. 'But exactly one grain is lacking!
550. You have stolen one grain!'
551. 'Not at all, not at all.'
552. He goes to him.
553. He looks all over him, his mouth, inside his nostril, everywhere.
554. There all over he looks, inside his ear.
555. 'Did you find what you were looking for?'
556. He had inserted the corn seed inside his little penis.
557. His two grandsons now arrive also.
558. He takes it out for him.
559. 'Here is your requested corn seed.
560. Now plant it.'
561. They return,
562. and they plant it also.
563. They invite their grandfather again.
564. 'Drink some corn drink Gramps.'
565. 'It must be as I thought!
566. That Grasshopper stole corn for you also.'
567. 'Of course.

568. Otherwise, what would our posterity eat?´
569. ´O.K.
570. Let it be so.
571. But it won´t grow quickly for them anymore.
572. It´ll take three months for it to grow for them.

Scene XXVIII: The Twins obtain the correct name for pifayo

---

573. After a while they invite him for pifayo drink also.
574. ´It must be that you stole the pifayo seed.´
575. ´Of course.
576. Otherwise what will our posterity have to drink?´
577. ´O.K. So be it.
578. But it will not grow quickly for them anymore.
579. It´ll take one year for it to give fruit for them.
580. He also wants its name now.
581. He asks the one living with him,
582. ´What does he call this?´
583. He does not say ´pʉ́ʉ́riy´ yet.
584. He just says its name is ´ta̧a̧churá.´
585. ´That´s not what it´s called!´
586. He says its real name.
587. He (Placenta) says to one of his servants,
588. ´How does he say its name?´
589. He is afraid of telling it.
590. ´Tell him only half of it.´
591. Of course he speaks its name loudly.
592. ´Pʉ́ʉ́riy,´ he says it.
593. ´Pʉ́ʉ́,´ his mouth twists then.
594. His mouth becomes little then.
595. ´What is it?´
596. ´Tííy,´ he can´t speak anymore.
597. Now he just babbles, ´si, si, si.´
598. ´Tíjɨ,´ he figures out its name.
599. ´Is it "pʉ́ʉ́ríy"?
600. Is that what he just said?´
601. ´Tííy,´ he can´t answer anymore.
602. He has been transformed into a flounder.

Notes

1. Yagua is the only extant member of the Peba-Yagua language family of northeastern Perú (Loukotka 1968). It is currently spoken as a first language by about 2,000 individuals. See T. Payne (1985) and D. Payne (1985) for details of Yagua grammar. See Powlison (1969) for ethnographic and folkloric information.

Abbreviations employed in the Yagua examples and in the text are the following:

1SG = First person singular
1DL = First person dual
1PLINC = First person plural inclusive
1PLEX = First person plural exclusive
2SG = Second person singular
2DL = Second person dual
2PL = Second person plural
3COL = Third person collective (same form as 3SG)
3SG = Third person singular
3DL = Third person dual
3PL = Third person plural
4 = Fourth person (1st paradigm only)
ARR1 = Arrival 1 (action takes place as subject arrives on current scene--no scene change implied)
ARR2 = Arrival 2 (action takes place as subject arrives on new scene--scene change implied)
AUX = Auxiliary
CL = Classifier
COP = Copula
DAT = Dative
DAY = Discourse particle
DEMO = Demonstrative
DIM = Diminutive
DIST = Distributive
E = Enclitic
HC = Head coding
INTS = Intensifier
INAN = Inanimate
IRR = Irrealis auxiliary (also referred to as AUX:IRR)
JIITA = Discourse particle
LOC = Locative (to, at, in)
MALF = Malefactive auxiliary (also referred to as AUX:MALF)
NIY = Semantically 'empty' morpheme following a left-dislocated pronoun
NOM = Nominalizer
NP = Noun phrase
PAST1 = 1st past tense (Action occurred a few weeks ago)
PAST2 = 2nd past tense (Action occurred a few months ago)
PAST3 = 3rd past tense (distant past)
PNP = Pre-predicate or pre-head noun phrase
Post-NP = Post-posed noun phrase

PROX1 = 1st degree of proximity (When attached to verbs, indicates that action occurred earlier in the day of speaking. When attached to other elements, including predicate nominals, indicates that action occurred in a location close to the speaker.)
PROX2 = 2nd degree of proximity (On verbs indicates action occurred a day or so ago; on other elements indicates action occurred farther away from speaker.)
RD = Referential distance
REP = Repetition (enclitic meaning ´again´ or ´also´)
SW = Sound word (idiophones, or onomotopoeic expressions)
TRNS = Transitivizer
VC = Verb coding

I would like to express my appreciation first of all to the Yagua people, especially Mamerto Macahuachi, Pedro Díaz, and Hilario Peña, all of whom contributed much to the preparation of this paper. Secondly, I would like to thank Sandra Thompson, Jack DuBois, and Doris Payne, all of whom read and commented on various drafts of this paper.

2. Forms from the paradigm represented in Chart A will always appear as prefixes. Therefore in the gloss line of examples, any gloss such as ´3SG´, ´1PLINC´, etc., which glosses a prefix refers to a form from this paradigm. On verbs such prefixes always code the subject (A or S), on nouns the same prefixes code possessors, and on postpositions the same prefixes code the object of the postposition. Occasionally a pronoun will appear with an enclitic attached. In order to distinguish such pronouns from the forms illustrated in Chart A, the gloss ´PRO´ will always accompany person and number specification of pronouns, e.g., ´3SG:PRO´. I will simply rely on linear order to distinguish the forms in Chart A from those in Chart B — forms appearing as prefixes are from Chart A, those appearing as enclitics are from Chart B.

3. The form **vurya** for the first person plural/dual inclusive is the allomorph regularly used for Class II, III, and IV stems.

4. In Yagua, placentas and uteruses are almost always expressed as possessed noun phrases. Contrary to what may seem natural from an English point of view, the possessors of these items are the children that are born with the placenta, or which come from the uterus. To refer to a person´s uterus is to refer to the uterus from which that person was born, even if that person is a full grown woman, and even if she happens to be pregnant. To refer to the uterus that is located inside a woman´s body, one must refer to her child´s uterus, even if she has never had a child and/or is not currently pregnant. In the text from which example (7) is taken, the placenta transforms into a human being and becomes a central character. As such, the term **munáá** ´placenta´ comes to be used as a proper name, and is therefore not required to be possessed. However, whenever it is possessed, it is possessed by Placenta´s elder brother, i.e., the child that was born with the placenta.

5. Full NPs referring to intransitive subjects do occasionally occur post-verbally without a coreferential prefix on the verb. However, the use of this coding device has much in common with ´S_o´ coding (see Sect.

2.2.). Quite independently from their defining characteristics based on subject coding devices, clauses which employ S_o coding have many features of non-verbal predications, as demonstrated in Sect. 2.2.1.

6. It is a general principle of Yagua discourse that one avoids the use of fully-specified noun phrases as much as possible, allowing the rich participant coding system to keep participants sorted out. Only in marked contexts, or to avoid ambiguity, are fully-specified noun phrases used. One strategy for avoiding the use of fully-specified noun phrases is to treat one plural participant as singular when two groups are interacting. In such cases it is the most topical group that is treated properly in terms of its semantic plurality, while the other group is treated as singular. For example, if adults and children are interacting, the adults will be coded as plural, while the children singular. If humans are interacting with animals, the humans will be plural and the animals singular. If ´good guys´ are interacting with ´bad guys´ (as is often the case in folkloric history narratives) it is predictably the ´good guys´ which are treated as plural while the ´bad guys´ are treated as singular:

(1) Rityęęryá rumusíy váriy,
riy-tąąrya rumusiy váriy
3PL-return from:there then
´They (good guys) returned from there,

sasiityatítyiiyanuntiryíy.
sa-siiy-ta-títyiiy-janu-ntiy-riy
3SG-run-TRNS-going-PAST3-REP-3PL
they (savages) chasing them again.´

In this example the savages are treated as singular in the second clause even though they are obviously a group of people and in other examples in the same story are treated as plural. The people being attacked (the clan to which the narrator belongs) are treated as plural. The ´good guys´ as a group are never treated as singular. This bending of the categories plural and singular is a very obvious feature of Yagua narrative, and is clearly used in order to avoid the use of fully-specified noun phrases.

7. Anaphoric zeros, as well as all other participant coding devices, are considered to constitute ´mentions´ of the participants they code. The terms ´refer to´ and ´mention´ are to be taken as equivalent in this paper. Essentially, high topicality for Givón is correlated with large number of mentions within a thematic paragraph, while low topicality is correlated with small number of mentions. It stands to reason that the participants that a text is ´about´ will be mentioned more often than others. This conception of topicality is logically independent of the notion of topicworthiness as discussed in Sect. 3.1.1, though the two notions are intimately related.

8. The assumption that participants that have not been mentioned in the discourse have no topicality whatsoever is extreme, but is a logical presupposition of the topic continuity framework. In Sect. 3.1 I present some of the problems inherent in this notion of topicality as number of mentions. In Sect. 3.3 I outline some of my crude attempts to

deal with these problems in my own topic continuity study of Yagua. The fact is that speakers are more likely to talk about certain entities apart from whether those entities have already been mentioned in the current text or not. It would be unfair to suggest that Givón does not recognize this fact, as more recent work has shown that he and his students are sensitive to other factors affecting topicality (e.g., Givón 1985). However, these factors have as yet not been incorporated into a quantitative study simply because they are so difficult to deal with. The best we have been able to do so far is to control for them.

9. Though see Sect. 4.2.7 for some evidence that this assumption may be mistaken.

10. Clause numbers are only approximate for several reasons. First, discourse units which clearly fill a breath group or conversational turn are counted as clauses even though they may not constitute a clause in the syntactic sense, e.g., 'Really?' (TC44), 'With just a blowgun' (TC53), 'Another day the same thing again' (TC102, 103) etc. Second, there were occasional misnumberings of clauses in the original transcriptions such that some numbers occurred twice and others were omitted. Hence, it will be noted that occasionally an annotated clause number occurs, e.g., TC358a, indicating that in the original transcription two distinct clauses received the number 358. This convention never results in a discrepancy of more than one or two clauses, and does not affect the referential distance counts at all, since the clause numbering system was not utilized in calculating the RD figures.

11. Although the claim made in Givón (1983c:19) refers specifically to 'rigid word order languages' such as English (SVO) and Japanese (SOV), the findings for Ute (Givón 1983d:196), a relatively 'free' word order language, yield the same conclusions. Further studies of Papago (D. Payne 1984) and Cayuga (Mithun 1984) also confirm this fact. This is not to say that Yagua is acting 'exactly like' Ute, Papago, Cayuga, Japanese, or English in this regard. In fact there are some very significant differences. For example, there is the fact that post-verbal subjects in Yagua require the use of a coreferential prefix on the verb whereas preverbal subjects (within the same intonation unit as the verb) preclude the prefix. Second, the preverbal NP device in Ute is used much more often than the post-verbal NP device (39 to 25), whereas in Yagua PNP is much less common than VC+NP (15 to 93). Finally, the RD indices for preverbal and post-verbal NPs in Ute differentiate much more strongly (10.84 for preverbal NPs and 1.48 for post-verbal NPs) than do the RD indices for PNP and VC+NP in Yagua (11.07 for PNP and 7.95 for VC+NP). Thus I conclude that the fact that the preverbal NP devices in both languages are more discontinuous than the post-verbal NP devices is a relatively minor similarity. In almost every other respect they are different.

12. The terms introduced here in quotes are impressionistically defined cover terms, not technical terms. The same is true of several other terms used in this section and elsewhere, such as 'recently,' 'highly topical,' etc. At this point in the state of the art in topic continuity studies, there is no standard by which we can evaluate in fine detail the differences between various values for the continuity

indices. Therefore, any conclusions regarding the significance of the difference between any two figures are purely impressionistic. The difference between an RD of 11.07 and 1.68 'feels' very significant, whereas the difference between 1.68 and 2.2 is less clear. I don't claim to have solved all the problems of the topic continuity framework in this paper. Hopefully, further research will be able to render the methodology more exact.

13. Ambiguity proved to be a very difficult measure to calculate due to the slipperiness of the notion of 'semantically compatible.' In the calculation of ambiguity for the 113 long-range devices used to code recently mentioned participants, I tried to be as conservative as possible in positing the presense of a semantically compatible referent on the discourse stage, precisely because I was interested in what factors other than strict ambiguity might be triggering the use of a stronger coding device. That is to say, some of the other six conditions I found for using extra-strong coding devices may well be subsumed under ambiguity in a framework that only recognized ambiguity as a possible reason for using a stronger than expected coding device.

References

Chafe, Wallace L. 1985. Cognitive constraints on information flow. Cognitive science program, University of California at Berkeley.

Clancy, Patricia M. 1980. Referential choice in English and Japanese narrative. The pear stories, ed. by Wallace L. Chafe, 127-202. New York: Academic Press.

Derbyshire, Desmond C. To appear. Topic continuity in Hixcariana. Native South American discourse, ed. by Joel Sherzer and Greg Urban.

Dixon, R. M. W. 1979. Ergativity. Lg. 55. 59-138.

Dorson, Richard M. 1975. Folktales told around the world. Chicago: University of Chicago Press.

Du Bois, John W. 1980. Beyond definiteness: the trace of identity in discourse. The pear stories, ed. by Wallace L. Chafe, 203-74. Norwood, NJ: Ablex.

Du Bois, John W. 1981. Ergativity and preferred argument structure in Sacapultec discourse, a preliminary outline. MS.

Fox, Barbara A. 1984. Discourse structure and anaphora in written and conversational English. University of California, Los Angeles doctoral dissertation.

Givón, T. 1983a. Topic continuity in discourse: the functional domain of switch reference. Switch reference and universal grammar, ed. by John Haiman and Pamela Munro, 51-82. Amsterdam: John Benjamins.

Givón, T. (ed.). 1983b. Topic continuity in discourse: a quantitative cross-language study. Amsterdam: John Benjamins.

Givón, T. 1983c. Topic continuity in discourse: an introduction. Topic continuity in discourse: a quantitative cross-language study, ed. by T. Givón, 1-41. Amsterdam: John Benjamins.

Givón, T. 1983d. Topic continuity and word order pragmatics in Ute. Topic continuity in discourse: a quantitative cross-language study, ed. by T. Givón, 141-214. Amsterdam: John Benjamins.

Grimes, Joseph. 1975. The thread of discourse. The Hague: Mouton.

Halliday, M. A. K. 1967. Notes on transitivity and theme, Part 2 Journal of Linguistics 3. 199-244.

Halliday, M. A. K. Ruqaiya Hasan. 1976. Cohesion in English. London: Longman.

Hopper, Paul J., and Sandra A. Thompson. 1980. Transitivity in grammar and discourse. Lg. 56. 251-99.

Hopper, Paul J., and Sandra A. Thompson. 1984 The discourse basis for lexical categories in universal grammar. Lg. 60.

Jaggar, Phillip. 1985. Referential choice in Hausa narrative. University of California, Los Angeles doctoral dissertation.

Klimov, G. A. 1977. Tipologia jazykov aktivnogo stroja. Moscow: Nauka.

Loukotka, Cestmir. 1968. Handbook of South American Indian languages. Los Angeles: University of California, Latin American Studies Center.

Mann, William C., and Sandra A. Thompson. 1983. Relational propositions in discourse. MS. To appear in discourse processes.

Mithun, Marianne. 1984. Is basic word order universal? Paper presented at the Symposium on Discourse Relations and Cognitive Units, June 1984, Eugene, Oregon.

Payne, Doris L. 1985. Aspects of the grammar of Yagua: a typological perspective. University of California, Los Angeles doctoral dissertation.

Payne, Thomas E. 1983a. Yagua object clitics: syntactic and phonological misalignment and another possible source of ergativity. The interplay of phonology, morphology, and syntax, ed. by John F. Richardson, Mitchell Marks, and Amy Chukerman, 173-84. Chicago: University of Chicago Press.

Payne, Thomas E. 1983b. Subject inflection of Yagua verbs. MS, available from the author.

Payne, Thomas E. 1985. Participant coding in Yagua discourse. University of California, Los Angeles doctoral dissertation.

Powlison, Paul S. 1969. Yagua mythology and its epic tendencies. Unpublished University of Indiana doctoral dissertation. Ann Arbor: University Microfilms.

Reinhart, Tanya. 1982. Pragmatics and linguistics: an analysis of sentence topics. Bloomington: Indiana University Linguistics Club.

Rumelhart, David E. 1975. Notes on a schema for stories. Representation and understanding: studies in cognitive science, ed. by Daniel G. Bobrow and Allan Collins, 211-36. New York: Academic Press.

Sacks, H., Emanuel Schegloff, and G. Jefferson. 1974. A simplest systematics for the organization of turn-taking for conversation. Lg. 50. 696-735.

Scancarelli, Janine. 1984. Referential strategies in Chamorro narratives. University of California, Los Angeles master's thesis.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. Grammatical categories in Australian languages, ed. by R. M. W. Dixon. Linguistic Series 22. 112-71. Canberra: Australian Institute of Aboriginal Studies.

A NOTE ON ERGATIVITY, S´, AND S´´ IN KARITIANA*

Daniel Everett

1 Introduction
2 Word order and absolutive marking
  2.1 Karitiana phrase structure
  2.2 Affirmative particles
  2.3 Pronouns
3 WH-questions and topics
4 An alternative analysis
5 On the default affirmative
6 Conclusion

1 Introduction

Ergativity has long been an important topic in linguistic research, from both a typological (Comrie 1978, Dixon 1979, Plank 1979) as well as a theoretical perspective (Levin 1983, Marantz 1984). The purpose of the present study is to contribute to the discussion of this phenomenon via a study of a fragment of the grammar of Karitiana, an Amazon language, focusing especially on its ergative-absolutive marking of affirmative particles and personal pronouns. The special features of the Karitiana (henceforth **K**) system are: (1) only absolutive marking need be stated by rule, ergative case being a default mechanism, and (2) K´s ergative marking offers interesting evidence in favor of the distinction between topicalization and **WH**-movement, along the lines first proposed by Chomsky (1977).

The discussion is organized as follows: first, we outline K´s basic phrase structure. Next, the ergativity facts are presented, along with the original analysis of these facts suggested in Landin (1980), according to which an apparently enigmatic asymmetry exists between **WH**-movement and topicalization. This is followed by an alternative account in which the facts fall out quite naturally as a consequence of the distinction between S´ and S´´.

2 Word order and absolutive marking

2.1 Karitiana phrase structure

According to Landin (1980), the basic word order in K is Subject-Verb-Direct Object:

(1)

sara	ty	nãka	-y	-t	taso	aka
alligator	big	affirmative	-eat	-tense	man	that
S		(erg)	V		O	

'The big alligator ate that man.'

(2)

õmãky	nã	-oko	-t	moroja
jaguar	affirmative	-bite	-tense	snake
S	(erg)	V		O

'The jaguar bit the snake.'

These data are interpreted in the work cited in support of the S-V-O order as basic, since they are unambiguous. As we see below, however, the interpretation of these sentences is in fact guaranteed by the affirmative particle rather than word order, thus removing the force of this argument. Noun phrase structure presents a somewhat clearer picture, however, the constituent order clearly being genitive (possessor) - head (possessed):

(3)

yjja	nãka	-y	-j	yj	pikkõm	pisyp
we	affirmative	-eat	-tense	our	monkey	meat
				GENITIVE	HEAD	

'We will eat our monkey meat.'

(4)

opok	nã	-tot	-ø	i	o
Indian	affirmative	-remove	-tense	3	head
	(erg)			GENITIVE	HEAD

'The Indian removed his head.'

In adpositional phrases, the head again is phrase final:[1]

(5)

owa	na	-ate	-tysot	pikkõm	i
child	affirmative	-pull	-aspect	monkey	3
	(erg)				

sypojo-sok	
tail	-on
NP	POSTPOSITION

'The child pulled the monkey by the tail.

(6)

nãka	-tat	-o	ga	-p
affirmative	-go	-tense	field	-to
(erg)			NP	POSTPOSITION

'He went to the field.'

2.2 Affirmative particles

As is seen in the above examples, K sentences manifest particles which Landin (1980) refers to as affirmative particles. According to Landin, the distribution of these particles is as follows:[2] **na-** is found following transitive subjects, while **ta-** follows intransitive subjects and transitive objects; **-ka-** is added to either form when preceding verb roots with initial stress.

(7) a. iso nāka -´y -t saryt kerep Ohey
fire affirmative eat -tense hearsay long ago name
´The fire ate Ohey long ago.´

b. *iso $\left\{\begin{array}{l}\text{nā-}\\ \text{ta(ka)-}\end{array}\right\}$´y -t saryt kerep Ohey

(8) a. y taka -tar -i
affirmative -go -tense
(abs)
´I will go.´

b. *y $\left\{\begin{array}{l}\text{ta-}\\ \text{nā(ka)-}\end{array}\right\}$ -tar -i

(9) a. ȳn nā -oky -j sojja
I affirmative -kill -tense pig
(erg)
´I will kill a pig.´

b. *ȳn $\left\{\begin{array}{l}\text{nāka-}\\ \text{ta(ka)-}\end{array}\right\}$oky -j sojja

(10) a. y ta -oty -j
I affirmative -bathe -tense
(abs)
´I will bathe.´

b. *y $\left\{\begin{array}{l}\text{taka-}\\ \text{nā(ka)-}\end{array}\right\}$-oty -j

We consider an alternative analysis of these facts in Sect. 4 below. However, we first need to consider ergative marking in K´s pronominal system.

2.3 Pronouns

K´s pronoun system is summarized in Table 1 (from Landin 1980:11):

TABLE 1
Pronominal System of Karitiana

<table>
<tr><th></th><th></th><th>1S</th><th>2S</th><th>3S+P</th><th>1P INCL</th><th>1P EXCL</th><th>2P</th></tr>
<tr><td>Subj. + Intr.
Obj. + Tr.
=Absolutive</td><td>Pre-vb</td><td>y</td><td>a</td><td>Ø</td><td>yj</td><td rowspan="2">yta</td><td>aj</td></tr>
<tr><td>Obj. +Tr.</td><td>Post-vb</td><td rowspan="2">ȳn</td><td rowspan="2">ān</td><td>i</td><td>yjja</td><td>ajja</td></tr>
<tr><td>Subj. + Tr.
=Ergative</td><td>Pre-vb
Post-vb</td><td colspan="4">Ø</td></tr>
</table>

The crucial features of this system for our present concerns are the preverbal forms of the first and second person singular forms. Landin (1980) explains their distribution by claiming that **ȳn** and **ān** are the subject forms for transitive verbs while **y** and **a** appear as subjects of intransitive verbs or as objects of transitive verbs, as in (11)-(13):

(11)

y	taka	-tar	-i
1	affirmative	-go	-tense
(abs)	(abs)		

´I will go.´

(12)

ȳn	a	ta	-oky	-j
1	2	affirmative	-kill	-tense
(erg)	(abs)	(abs)		

´I will kill you.´

(13)

ān	y	ta	-oky	-j
2	1	affirmative	-kill	-tense
(erg)	(abs)	(abs)		

´You will kill me.´

Having completed our brief overview of the relevant features of K syntax, we turn now to consider an interesting contrast in ergativity marking between topicalized structures and **WH**-questions. We argue in Sect. 4 that these facts are strongly supportive of the analysis of topicalization and wh-movement developed by Chomsky (1977; 1981).

3 WH-questions and topics

Landin (1980:15) claims that topicalization in Karitiana is produced by a movement rule, fronting the topicalized constituent, as in (14):

(14)	a.	y	taka	-tar	-i	se -pip
		1	affirmative	-go	-tense	water-into
		(abs)	(abs)			

'I will go into the water.' (normal declarative)

b. se -pip y taka -tar -i ___
'Into the water I will go.' (topicalized reading)

A slight difference arises, however, when the topicalized element is the (underlying) direct object, as in (15):

(15)	a.	yjja	na	-pyn	mora
		1 pl	affirmative	-kick	ball

'We kick the ball.'

	b.	mora	yjja	ti	-pyn ___
		ball	1 pl	topi	-kick

'The ball, we kick (it)'

(15) differs from cases of nonobject topicalization, as in (14), in the appearance of **ti** 'topic' on the verb. Landin (1980) claims that the purpose of **ti** in these structures is to prevent ambiguity, signalling that the first NP in the clause is to be interpreted as topic. **WH**-questions are formed in a similar fashion.

(16)	mõramõn	a	ti	-pa	-tynh
	what	2	topic	-weave	-aspect

'What are you weaving?'

According to Landin, (16) may be derived from an underlying structure along the lines of (17).[3]

(17) ___ a ti -pa -tynh mõramõn

Landin goes on to observe that topics and **WH**-questions might be analyzed as arising from the same rule of "front topic/**WH**-word, insert **ti**." However, as he correctly notes, such an analysis is not complete in light of examples such as (18):

(18) seppa ỹn ti -pa -tynh
basket 1 topic -weave -aspect
(erg)
´A basket, I am weaving.´

ÑppThe problem, of course, is why the pronoun in (17) should be absolutive while the pronoun in (18) is ergative. Both bear the same thematic role, and both follow the item associated with the patient role. It cannot be claimed that interrogatives are inherently absolutive in K due to **WH**-questions of adjuncts, as in (19):[4]

(19) mõrasõg ãn i pa -tynh seppa
why 2 3 weave -aspect basket
´Why are you weaving a basket?´

Thus, the fact that the second person pronoun **a** in (16) is marked absolutive cannot be explained solely on the basis of its being in the interrogative mood. Landin (1980:27ff.) concludes his study of ergativity by claiming:

> ...In Karitiana it is not the transitivity of the verb which determines the ergativity value of the accompanying pronouns, but rather the number of associated NPs. If a transitive verb has a subject and no object, then the subject pronoun associated with the verb will be absolutive, and not ergative.
>
> The curious and unexplained fact is that a wh-word like **mõramõn** does not count as an object NP, while a pronoun or even an NP that has been removed by topicalization does count."

In the remainder of our discussion, we show that Landin´s conclusion is in fact incorrect, and that the ergativity facts, properly analyzed, fall out as a direct consequence of the S´ and S´´ nodes proposed in Chomsky (1977,1981).

4 An alternative analysis

We begin this section with a summary of the main points of the analysis of ergativity in K given above.

(20) Ergative Marking (preliminary version):

a. The affirmative particle is absolutive following an intransitive subject or transitive object; otherwise it is ergative.

b. First and second person singular preverbal pronouns are ergative when both subject and direct object are overtly present; otherwise they are absolutive.

Note that (a) and (b) are quite different. Whereas (a) is stated in terms of a preceding argument, (b) is stated in terms of the total number of overt arguments in the clause, regardless of their position relative to the pronoun. A reasonable move would be to attempt to collapse (20a) and (20b) into a single statement. It seems difficult to improve on (20a), however, in view of sentences such as (21) (from R. Landin 1982:4):

(21) Ohey taka -´y saryt kerep Isoason
name absolutive eat hearsay long:ago name
´Isoason ate Ohey long ago.´[5]

In (21), where the order is O-V-S (cf. note 5), no ambiguity arises with regard to the grammatical relation borne by a specific NP, since the presence of the absolutive **taka-** following **Ohey** clearly marks **Ohey** as the direct object (cf. (22)):

(22) Ohey nāka - ´y saryt kerep Isoason
´Ohey ate Isoason long ago.´

On the other hand, we might attempt to restate (20b) as (23):

(23) Ergative Marking (second version):

Mark first and second person singular preverbal pronouns as absolutive when they are intransitive subjects or direct objects or they follow the direct object or intransitive subject; otherwise, mark them ergative.

Note that (23) will account for all the facts, including the appearance of the absolutive **a** ´2 person singular´ in (17), as opposed to the ergative **ān**, which is predicted by (20b). This is so if we assume, contrary to Landin (1982), that the **WH**-word **does** count as an argument. Then in (17), the pronoun follows the direct object and is absolutive, as predicted by (23). The question remains, however, as to why in (18) the pronoun is ergative, since it immediately follows the direct object. Our suggestion is to limit (20a) and (23) to **S´** as in (24):

(24) **Ergative Marking** (final version):

> Mark the **designated element** (= affirmative particle, first and second singular preverbal pronouns) as absolutive when it **is** the intransitive subject or direct object or is immediately **preceded by** the direct object or intransitive subject **in S´**; otherwise mark it ergative.[6]

(24) immediately explains the contrast between (16)-(18), given that topics are immediately dominated by **S´´**, whereas **WH**-questions involve a **WH-operator** in **COMP**, immediately dominated by **S´**, as in (25) and (26):

(25) **Topicalization:** $[_{S''}[_{Topic}\dots]\ [_{S'}\dots]]$

(26) **WH-Questions:** $[_{S''}\dots[_{S'}[_{COMP}\ \text{WH-operator}]\dots]]$

Therefore, we can understand K´s ergative marking system, as stated in (24), to be quite suppportive of the structural differences between **Topics** and **WH**-questions proposed in most work on the subject in generative theory.[7] This is true in spite of the fact that in current work on phrase structure in generative theory, the status of **S´´** is unclear. According to recent suggestions of Chomsky (class notes, p.c.), **S´** is a projection of the COMP node corresponding in **X´** terms to **COMP´´**. However, it is not clear what **S´´** would be a projection of in current terms. In a recent paper, Pullum (1985) suggests the following phrase structure rules:

(27) S´ ⟶ (TOP)S

(28) S´´⟶ COMP S´
where S´ is the initial symbol.

This seems unlikely to account for Karitiana ergativity, however, since then wh-words moved to COMP would be predicted to be less closely associated to S than Topic. As the facts above show, however, it is the Topic element which behaves as though it were less affected by sentence internal syntax. A way of capturing these relations in current theoretical terminology would be to assume that Topic and S´ (= **COMP´´**) have different underlying sources and that Topic is Chomsky-adjoined to S´ in the course of the derivation, as in (29):

(29) a. $[_{\alpha}\text{TOPIC}]\ [_{\text{COMP}''}\ \ldots]$

b. $[_{\text{COMP}''}\ [_{\alpha}\text{Topic}\ [_{\text{COMP}''}\ \ldots]]]$

Then ergative marking would apply to COMP'' prior to (29b). This is but one of several alternatives that come to mind for capturing Chomsky's (1977) S'/S'' distinction in the present model.

5 On the default affirmative

One question remains unanswered by the account just given, however. Namely, why should a language such as K have a default affirmative marker to begin with? That is, why should **nã(ka)-** even appear at all in those cases such as (30) where no argument precedes and the sentence is clearly intransitive (cf. (6) above)?

(30)	Ø	nã	-oty	-j
	3 (erg)	affirmative	-bathe	-tense
	'(He) will bathe.'			

ÑppThere is a simple answer to this question which may be seen by comparing (30) with (31):

(31)	Ø	oty
	he	bathe
	'He will **not** bathe.'	

In other words, without the affirmative marker, a sentence is interpreted as negative (tense is also obligatorily absent). As Landin (1980:19) notes, this violates certain proposed universals of negation, in which it is claimed that negative clauses should be morphologically more complex than affirmative clauses (cf. Dahl 1978, Payne 1978).

Due to this state of affairs, however, the affirmative particle is necessary to give the right reading, explaining how the particle **nã(ka)-** could take on a default marking. Thus, in a sense, absolutive marking is the rule-governed case, with ergative marking following under the "elsewhere" condition of the rule.[8]

6 Conclusion

In this brief note, we have seen that ergative case marking on preverbal pronouns and the affirmative marker is stateable in terms of a single generalization (cf. (24)).

Further, in that this marking only takes into account constituents of S', not considering S'' constituents such as **Topic**, it offers support for the distinction between topicalized and interrogative structures proposed in Chomsky (1977, 1981).

Notes

* Karitiana is a member of the Arikem family of the Tupi phylum and is spoken by approximately eighty people near Porto Velho, in the state of Rondonia, Brazil. The orthography used here was developed by David and Rachel Landin. Aside from **y**, used to represent /i/, and ´ for /?/, all other symbols are straightforward.

We would like to thank the Landins for their work on this language, without which the present paper would obviously be nonexistent. Also thanks to Noam Chomsky, Ken Hale, Dave Landin, and Geoff Pullum for comments on earlier versions of this paper.

The research was supported in part by the American Council of Learned Societies and Grant BNS-8405996 from the National Science Foundation.

1. It is not clear from the data presented, however, whether K even has adpositional elements in the traditional sense. A plausible case could be made that such "postpositions" in K are simply case markers. We will not take up this question here, however, since it has little bearing on the central issue.

2. It is not clear to us whether these morphemes are best analyzed as independent particles, verbal prefixes, or second position clitics. Since this is not crucial for our present purposes, we will represent these as prefixes, as per Landin (1980).

3. We return directly to the problem of why (16) and (17) should be marked absolutive.

4. The affirmative particles **ta(ka)-** and **nã(ka)-** are never found in interrogatives or (direct object) topicalized structures.

5. (21) wreaks havoc with the claim that K´s basic word order is S-V-O. Note that (21) cannot be analyzed as a topicalized structure, due to the absence of the topic marker, **ti-**. Therefore, the argument used by Landin (1980) in favor of S-V-O, namely, the lack of ambiguity in such structures, is vitiated, since it is the ergative or absolutive marking on the following affirmative particle which guarantees correct interpretation of the relevant grammatical relations, not word order. In fact, sentences like (21), very common in R. Landin (1982), make K appear similar to Tupi languages in general in manifesting free word order (cf. Harrison (to appear)).

6. In (24), **subject** and **direct object** refer to logical relations and not surface positions. Also, we suppose that if K word order is in fact free at S-level (although pragmatically constrained via discourse considerations) then all structures are generated as seen on the surface, with no movement.

7. Whether or not topicalization involves raising of a null **WH**-operator from object position in examples such as (18) is irrelevant since our statement in (24) is to be construed as applying to **overt** arguments only. The same applies for (i) where, according to Landin (1980:9) the affirmative **nã-** is immediately preceded by a **null** third person pronoun.

(i) ỹn Ø nã -oky -j i
1 3 affirmative -kill -tense 3
(erg) (erg)
'I will kill him.'

Similarly, in (30) below, where the verb is intransitive, the affirmative is still ergative, offering stronger evidence yet in favor of the characterization of ergative case as a default marking.

8. It is tempting to speculate that pronouns came to be marked ergative by default rather than by their semantic roles as a consequence of the conditions imposed on the rule by the affirmative particle. Thus, whenever the affirmative particle became obligatory, the entire rule (24) was modified to handle it, pronouns being affected due to the application of the rule across the board.

References

Chomsky, N. 1977. On WH-movement. In Formal Syntax, ed. P. Culivover, T. Wasow, and A. Akmajian. New York: Academic Press.

Chomsky, N. 1981. Lectures on government and binding. Dordrecht: Foris.

Comrie, B. 1978. Ergativity. In Syntactic Typology, ed. W. Lehmann. Austin: University of Texas Press.

Dahl, O. 1978. Typology of sentence negation. Linguistics 17.

Dixon, R.M.W. 1979. Ergativity. Language 55.

Harrison, C. to appear. Ergativity and typological disharmony in Guajajara. In Handbook of Amazonian Languages, ed. D. Derbyshire and G. Pullum (eds.), Berlin: Mouton.

Landin, D. 1980. An outline of the syntactic structure of Karitiana sentences. Unpublished M.A. thesis, University College London.

Landin, R. 1982. Word order variation in Karitiana. Unpublished MS, Brasilia: Summer Institute of Linguistics.

Levin, B. 1983. On the nature of ergativity. Unpublished doctoral dissertation, MIT.

Marantz, A. 1984. On the nature of grammatical relations. Cambridge, MA: MIT Press.

Payne, J. 1978. Negation. To appear in Papers from the Language Typology and Syntactic Fieldwork Project.

Plank, F. (ed.) 1979. Ergativity. New York: Academic Press.

Pullum, G. 1985. Assuming some version of the X-bar theory. Mimeographed, Syntax Research Center, University of California at Santa Cruz.

SOME ASPECTS OF ZAPOTECAN CLAUSAL SYNTAX

Stephen A. Marlett

1 Introduction
2 Types of pronouns
  2.1 Overview of personal pronouns
  2.2 Reflexive and reciprocal pronouns
  2.3 Deictic pronouns
  2.4 Generic nouns
3 Semantic distinctions in pronouns
4 Number marking
5 Syntax
  5.1 Word order
  5.2 Distribution of free and bound pronouns
  5.3 Coreference situations
    5.3.1 Relative clauses
    5.3.2 Reflexive clauses
    5.3.3 Subject-possessed objects
    5.3.4 Parallel verbs
    5.3.5 Adverbial clauses
    5.3.6 Purpose clauses
    5.3.7 Complement clauses
    5.3.8 Discourse conditions
  5.4 Copying
    5.4.1 Fronted noun phrase
    5.4.2 Emphatic subject
    5.4.3 Modal verbs
  5.5 Other constructions where a noun phrase is missing
    5.5.1 Imperatives
    5.5.2 Locative subject
    5.5.3 Unspecified object
    5.5.4 Meteorological verbs
  5.6 Special clausal constructions
    5.6.1 Passive
    5.6.2 Comitative object
    5.6.3 Instrumental object
    5.6.4 Indirect object
    5.6.5 Oblique object
    5.6.6 Dative subject
    5.6.7 Locative as subject
    5.6.8 Causative clauses
    5.6.9 Noun incorporation
    5.6.10 Impersonal intransitive clauses

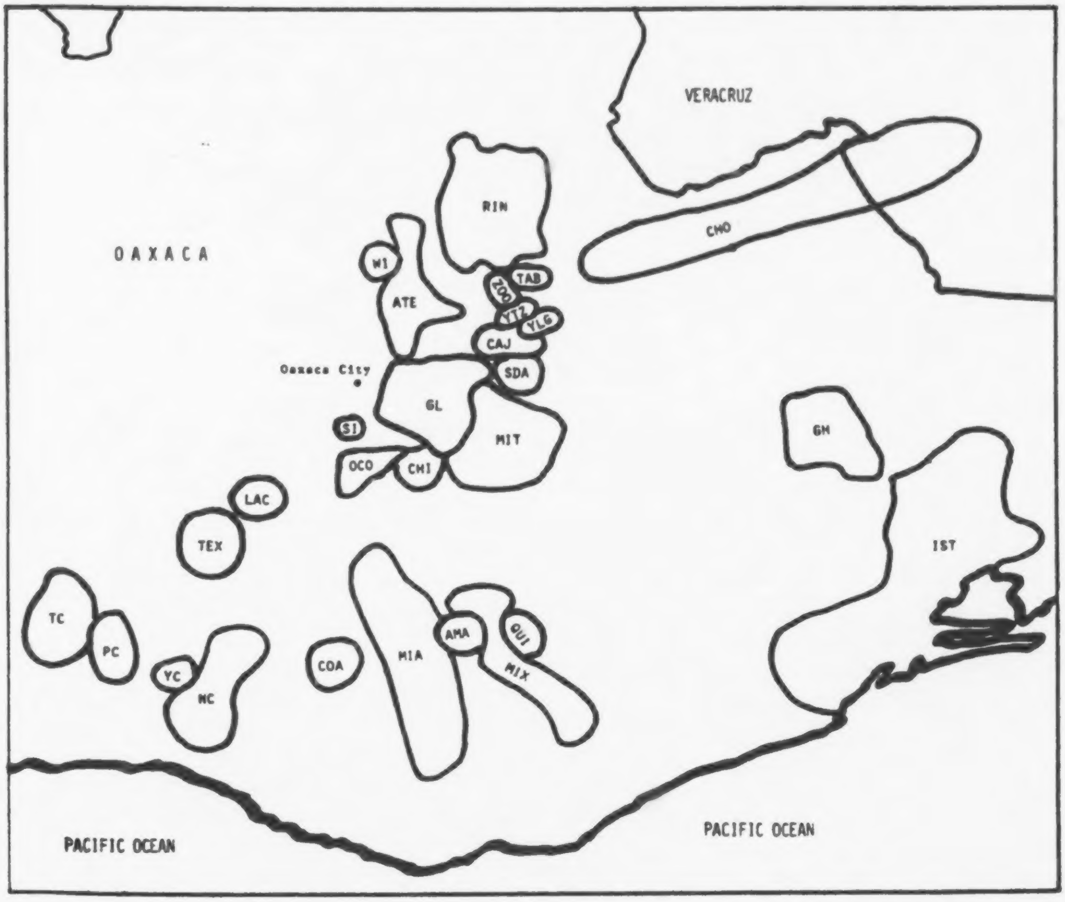
1 Introduction

In this paper I discuss pronouns and various syntactic constructions in Zapotecan (Z) languages. My major goal is to present a broad overview of the similarities and differences which exist among the various members of the Zapotecan group in these syntactic constructions, which I generally discuss in theory-neutral terms.

Zapotecan (see map) is generally considered as involving two branches--Chatino and Zapotecan proper (Suárez 1983:xvi). Of the latter there are numerous distinct variants (Rendón 1967). The actual number has been estimated to be as high as thirty-eight (Egland 1978). While this high figure has been contested--usually for political reasons and without substantial linguistic basis--the fact remains that Egland (1978) represents the only serious attempt to collect and analyze data relating to the mutual intelligibility of the speech of many villages. The Z languages are spoken by some 250,000 people who reside primarily in the state of Oaxaca, Mexico (Horcasitas de Barros and Crespo 1979). (A small percentage live in Veracruz.) The Chatino (Ch) languages, which are spoken by approximately 12,000 people in the state of Oaxaca, include five distinct groups (Egland 1978). I believe that the Zapotec-Chatino division is justified on various grounds, and therefore distinguish between these two major branches of Zapotecan repeatedly. I do not make any attempt here to group the various languages of the Zapotec branch.

The data are drawn primarily from fieldwork done by members of the Summer Institute of Linguistics in twenty-four variants of Zapotec and four variants of Chatino. I make no claim here as to the degree of differences between these variants, some of which are obviously much closer to each other than to others (see Egland 1978). In some cases it is well-known that a certain variant is extremely close to another. I include the data from all of these variants simply because the data are fairly reliable, in some cases published, and I was able to obtain further information from the source persons easily. These variants are referred to by the abbreviations in Appendix A. The published sources from which I obtained some data are listed below, but I cite data without explicit references. I gratefully acknowledge the large amount of information which individual Z linguists provided freely. Throughout this work I use a standardized orthography for the data below, the peculiarities of which are noted in Appendix C. I do not represent tone, which is nevertheless distinctive generally in Zapotecan.

Map showing general location of Zapotecan languages


VERACRUZ
OAXACA
RIN
CHO
MI
ATE
TAB
ZOO
YTZ
YLG
CAJ
Oaxaca City
SDA
GL
SI
MIT
OCO
CHI
GH
LAC
TEX
IST
TC
PC
YC
NC
COA
MIA
AMA
QUI
MIX
PACIFIC OCEAN
PACIFIC OCEAN

2 Types of pronouns

I categorize the pronouns of Z languages into the following four types: personal (with both free and bound variants), reflexive (and reciprocal), deictic, and a generic noun type. (I do not discuss interrogative or indefinite pronouns.) In this section I illustrate these types of pronouns. The semantic parameters of these forms are discussed in Sect. 3, and their syntax in Sect. 5. Pluralization is discussed in Sect. 4.

2.1 Overview of personal pronouns

A very common feature of Z languages is the existence of both "free" and "bound" personal pronouns. (See Appendix D.) Bound pronouns are phonologically dependent and must attach to another word, usually a preceding word, but sometimes a following morpheme. (Speck (1984) distinguishes between two phonological boundaries for these bound pronouns.) The bound pronouns are typically, but not always, distinct from the free pronoun. Even when they are the same, however, I write a hyphen before the pronoun when it occurs in a position which would require a bound pronoun. Other differences between free and bound pronouns are noted below. The distribution of free and bound pronouns in clauses is discussed in Sect. 5.2. Compare the following sentences from YTZ. (1a) contains a free pronoun as subject (a possible, but uncommon, construction), and (1b) a bound pronoun as subject.

```
(1) a. (YTZ)  gw-tas       le?
              COMPL-sleep  2s
              'You slept.'

    b. (YTZ)  gw-tas-o?
              COMPL-sleep-2s
              'You slept.'
```

Pronominal forms in Z languages generally do not vary according to grammatical function (in other words, they do not reflect case morphologically), although the distribution of free and bound forms is determined partially in most Z languages by grammatical function. (This important matter is discussed in Sect. 5.2.) One exception to this generalization is the first person free form in GL which is **na?a** generally, but **nare** when it is the object of the verb.

Sometimes a preposition may fuse with the bound pronoun, as in MIT, in slightly idiosyncratic ways. In the Ch languages, object pronouns are typically preceded by a certain particle (?ĩ). This particle takes the bound form of the pronoun after it. Since the corresponding form for first and second person singular sometimes involves a fusion of these elements, there is now a limited type of case marking in pronouns. These forms are labelled "genitive" below.

(2) (NC)

	Subject		Genitive
	Free	Bound	
1s	na?	(nasal.)	?ñã
2s	nu?wĩ	(tone)	?ĩ
1p inc	na	-na	?ĩ-na
1p exc	wa	-wa	?ĩ-wa
2p	?ũ	-ũ	?ĩ-ũ

I have not listed all of the allomorphy associated with bound pronouns in Appendix D. For example, in lieu of the third person inanimate form /-n/ in YTZ, the allomorph /-e/ occurs in the reflexive and reciprocal pronouns. Phonologically conditioned allomorphs are common. Also, most of these languages display some degree of dialectal variation within them which is simply ignored here.

I follow other analysts of Z in generally not analyzing the bound pronouns on verbs as agreement affixes for three reasons. First, they are in complementary distribution with the free pronouns and noun phrases in most cases (except under conditions discussed in Sect. 5.4). RIN is exceptional in this regard since the bound pronoun (subject or object) occurs on the verb regardless of the presence of a coreferential noun phrase. If there are two bound pronouns, they both occur on the verb. It is the only Z language known to have developed this "agreement" type of system. The word order in RIN is also different from other Z languages. Whereas the bound pronouns occur in the typical order VSO, the noun phrases occur in the order VOS in RIN. Two examples are given below.

(3) a. (RIN) guiX-ã? bɨni? ni?
COMPL.go-3resp man that
´That man went.´

b. (RIN) gu-gu?-ã-ba? bɨa? bɨni? ni?
COMPL-hit-3m-3anim horse man that
´That man hit the horse.´

Second, the bound pronouns are generally subject to the same word order constraints as free pronouns and noun phrases (see Sect. 5.1). Third, the bound pronouns may function in the plural marking system as noun phrases (see Sect. 4).

The first and second person free pronouns (both singular and plural) are monomorphemic forms in many Z languages, as shown for YTZ and NC in (4). The TEX and IST forms also given in (4) show that this is not always the case. But in several Z languages the first and second person plural forms are not monomorphemic, but involve a stem followed by a bound personal pronoun. The stem employed in IST is **la?a-** (which has no independent meaning). The plural morpheme **ka-** also optionally occurs in the first person plural inclusive form in IST. The plural TEX pronouns in (4) involve the plural morpheme **de**, and the first person plural inclusive form also involves the stem **ub** ´self´.

(4)	YTZ	NC	TEX	IST
1s	nada?	na?	yã	na?a
2s	le?	nu?wĩ	ru	li?i
1p inc	ǰio?o	na:	de-ub-na	la?a(-ka)-nu
1p exc	neto?	wa	dẽ-yã	la?a-du
2p	le?e	?ũ	de-ru	la?a-tu

Third person free pronouns, on the other hand, are generally morphologically complex forms (not in Ch, however). Plurality is indicated by a plural morpheme in some languages. In YTZ and IST, for example, the third person forms are composed of the stem **lle-** and **la?a-**, respectively, followed by a plural morpheme if the referent is plural, followed by a bound pronoun, as shown in (5). (I discuss plural marking in detail in Sect. 4.) The stem **la?a** (or its cognate) is used as a "focus" proclitic of some type with any subject noun phrase in some Z languages (e.g., ATE and MIX), but it does not have this use in many other languages.

(5)

	Third singular		Third plural	
	YTZ	IST	YTZ	IST
resp	ll-e?	la?a-be	lle-ga?ak-e?	la?a-ka-be
fam	lle-bo?	-------	lle-ga?ak-bo?	---------
anim	lle-b	la?a-me	lle-ga?ak-Eb	la?a-ka-me.
inan	lle-n	la?a-ni	lle-ga?ak-En	la?a-ka-ni

TEX has no equivalent stem, employing instead deictic pronouns which are morphologically complex in a similar way (see Sect. 2.3).

In some languages third person pronouns are commonly omitted in a running narrative, as discussed in Sect. 5, but in IST it can also be established that there is a "zero" third person pronoun since the pronominal stem and also the pluralizer can occur before it: **la?a-ka-0** (stem-PL-3). I have not listed this pronoun in Appendix D, however, since it does not contrast semantically with other third person pronouns.

The bound pronouns in some Z languages are recognizable as reduced forms of the free pronouns (if the latter are monomorphemic). Compare, for example, the first and second person pronouns of YTZ and GL in (6). (The parenthesized elements in the YTZ forms appear with some verbs and after certain morphemes; the parenthesized elements in the GL data are sometimes present, but are usually omitted.)

(6)

	YTZ		GL	
	Free	Bound	Free	Bound
1s	nada?	-(d)a?	na?a	-a?
2s	le?	-(d)o?	liu?	-[illegible]? (fam)
1p incl	ʒio?o	-ʒo	dunnɨ	-n(ɨ)
1pexcl	neto?	-to?	-----	----
2p	le?e	-le	la?a-t	-t(ɨ) (fam)

When a bound pronoun occurs on the verb in a Z language, it always follows the verb. (A bound pronoun is not necessarily suffixal elsewhere. It may be the first element of a deictic pronoun, as is shown below.)

In many Z languages (none of the Ch languages, however), more than one bound pronoun may be attached to the verb, under conditions made explicit in Sect. 5.2. In the majority of cases, a non-subject **bound** pronoun must be third person. Some examples are given below.

(7) a. (YTZ) b-neZ̈Xw-e?e-bo?o-b
COMPL-give-3resp-3fam-3anim
'He gave it to him.'

b. (TEX) b-za?ak-yu-y
COMPL-see-3m,fam-3m,fam
'He saw him.'

Z languages generally have a class of adjectives separate from the class of verbs. Adjectives, while generally not taking aspect marking, may be inflected for person using the bound pronouns, as in (8a). A copular verb may also be used, as in (8b-c).

(8) a. (YTZ) bitek gual-o?
how strong-2s
'How strong you are!'

b. (YTZ) Z̈oJ̈e n-akE-n
pretty STAT-be-3inan
'It is pretty.'

c. (IST) gu-ka-ni galan yanna ma ko?
COMPL-be-3inan pretty now -- no
'It was pretty, but not anymore.'

The pronoun attaches to the copula in (8b-c), but to the adjective in the following example from IST. I do not have much data on such sentences cross-linguistically and thus merely point out the facts.

(9) (IST) gu-ka nania?-ya?
COMPL-be hot-1s
'I am hot.'

In equative sentences which do not involve a copula, the free pronoun is used in IST but the bound pronoun in MIT. Again, I do not have much data on such sentences cross-linguistically.

(10) a. (IST) maystru la?a-be
teacher 3hum
´He is a teacher.´

b. (MIT) ?et stad-ti-tu-ni
NEG father-NEG-2p-3fam
´He is not your father.´

All Z languages which have bound pronouns also use them with nouns to indicate possessor with nouns (at least some group).

(11) a. (YTZ) yičX-a? ´my head´
head-1s

b. (GL) ş̌-libr-a ´my book´
POSS-book-1s

c. (TEX) bay-yu ´his waist´
waist-3m

d. (NC) kẽ ´my head´
head.1s

Some body parts function as prepositions in Z languages, and bound pronouns are used with these as well as some other prepositions, as in (12a-f). With other prepositions, however, free pronouns are used, as in (12g-j). The latter class of prepositions generally seem to be loanwords. (In TC free pronouns occur after all prepositions, however.)

(12) a. (YTZ) bi?in zeǰa kwit-o?
boy STAT.stand side-2s
´the boy who is standing beside you´

b. (YTZ) koŽE?E-bo? ´behind him´
back-3fam

c. (TEX) ri?iñ nẽ ´my work´
work of.1s

d. (TEX) b-rib-ã nañ-ma
COMPL-emerge-1s inside-3anim
´I emerged from it.´

e. (YTZ) bia če-bo? ´his animal´
animal of-3fam

f. (NC) ngusa ?ño 'my table'
table of.1s

g. (YTZ) par nada? 'for me'
for 1s

h. (YTZ) ka nada? 'than me'
than 1s

i. (YTZ) gwnn-e? fabor nada?
COMPL.speak-3resp favor 1s
'He spoke in my favor.'

Finally, bound pronouns commonly occur with quantifiers in Z languages.

(13) a. (GH) to-me 'one of them'
one-3hum

b. (YLG) čop-ak-be? 'two of them'
two-PL-3fam

c. (YLG) ball-en 'some of them'
some-3inan

2.2 Reflexive and reciprocal pronouns

Reflexive and reciprocal forms in many Z languages involve a special morpheme followed by a bound pronoun. (See Appendix E.) Note the following examples:

(14) **Reflexive:**	YTZ	TEX
myself	kwin-a?	ub-a
himself (fam)	kwim-bo?	ub-yu
themselves	kwin-ga?ak-bo?	
ourselves (incl)	kwin-čo	ub-na
Reciprocal:		
each other (3p)	lXwežX-e?	losa?-y
each other (1p incl)	lXwežX-čo	losa?-n

(Third person reciprocal pronouns in YTZ may optionally have the plural morpheme **ga?ak** preceding the bound pronoun.)

In the reflexive forms of some Z languages it is the free pronouns that are used. In IST, the reflexive stem precedes the free form. In CHI the reflexive stem intervenes between the pronoun stem and the bound pronoun if the free pronouns are morphologically complex.

(15)

	IST	CHI	MIX
myself	laka na?a	na?-gaka	leke na?a
himself	laka la?a-be	la?-gaka-ba	leke me
ourselves	laka la?a-nu	la?-gaka-nu?	

Some reciprocal forms have clear etymologies. YTZ **lXwe§X** is a noun meaning something like ´companion´ (Butler 1976), and **sa?** in OCO means ´another´.

In some languages there is no reflexive pronoun (strictly speaking) at all. I discuss the syntax of reflexive clauses in 5.3.2 and these languages are discussed there.

The reflexive pronoun may be used as an emphatic pronoun, as in the following examples:

(16) a. (YTZ) kwin-a? b-et-a? yin? be§-en?
self-1s COMPL-grind-1s chile tomato-the
´I myself ground the chili and tomato sauce.´

b. (YTZ) gw-yeX kwin-a?
COMPL-go self-1s
´I myself went.´

c. (PC) ykwi:? y?ni: tnyä kä?
self.2s COMPL.do.2s work there
´You are the one who did that work.´

In some Ch languages, the reflexive pronoun can be used in lieu of a pronoun.

(17) a. (TC) ?a nu?ü laka Leo // ykwĩ?
Q 2s CONT.be self.1s
´Is that you, Leo? // Yes, it´s me.´
(The answer is literally, ´Myself.´)

b. (TC) ?a nā ndi?ī ykwi?
Q NEG CONT.be.living self.3
´Isn´t the man of the house at home?´
(The question is literally, ´Isn´t himself living?´)

Some descriptions of Z claim that the reflexive or reciprocal morpheme is not a separate word which is followed by a bound or free pronoun (as in the YTZ examples above), but rather a verbal affix. In most cases it is hard to present strong arguments for one analysis over the other. The IST reciprocal morpheme is one example in which the verbal affix analysis is claimed.

(18) (IST) ka-gin̄e-sa?a-ka-be
CONT-hit-RECIP-PL-3hum
´They hit each other.´

(rather than * ka-gin̄e sa?a-ka-be)

I list the reciprocal forms in Appendix E without deciding between these alternative analyses. In all cases the reciprocal morpheme is followed by a bound pronoun or a noun phrase.

2.3 Deictic pronouns

Z languages typically distinguish three or four deictic elements. In some languages the same elements are used as adverbs in a sentence or as adjectives. Examples from TEX are given below.

(19) a. (TEX) gwa-y ze?
COMPL.go-3m there
´He went there.´

b. (TEX) b-yap-ā yu? i?i
COMPL-come-1s house here
´I came to this house.´

In other languages, such as MIX, the form of the deictic is different depending on its use. The forms listed below are those which occur as adjectives.

(20)

TEX	MIX	NC		
i?i	ri?i	re	´here´	
re	ki	kwa	´there´	(visible to speaker)
--	rek	kā?	´there´	(farther away)
ze?	ko	---	´there´	(not visible to speaker)

In TEX third person bound pronouns occur before these deictic particles to form deictic pronouns.

(21)

(TEX)	Bound Pronoun	´this one´	´that one´	´that one´
3 fam m (m.s.)	-yu	yui?i	yure	yuze?
3 fam f/inan (m.s.)	-n̄i	nii	nde	nde?
3 resp f/deity (m.s.)	-mi	mii?i	mire	mize?
3 anim	-ma	mai?i	mare	maze?

The deictic pronouns in TEX pattern syntactically like free pronouns, as in (22). As I noted in 2.1, TEX has no other third person free pronouns.

(22) (TEX) ornu bi b-dab-ma tub yu-ze?
when already COMPL-swallow-3anim one 3m-there
´Now after it (alligator) had swallowed one of them...´

In ATE any bound pronoun can occur before the deictic element to form deictic pronouns, as in (23).

(23) (ATE) tu?-ā? ´that one of ours´
1p,excl-there

The complementizer used in relative clauses in several languages may also be followed by a deictic to form a deictic pronoun.

(24) (NC) nu-re ´this one´
COMP-here

Appendix F summarizes information about morphologically-complex deictic pronouns.

2.4 Generic nouns

In this section I discuss a class of nouns which we call generic nouns (following Butler 1980). Generic nouns are a bit hard to define, however. In the clearest cases (as in YTZ) they are nouns meaning ´animal´, ´thing´, ´person´, etc. In less clear cases (as in NC) we also mean by this term certain shortened forms related historically to such nouns.

The animal pronoun in most, if not all, Z languages is either the same as or very similar to the noun for ´animal´. One might claim that these are all generic nouns in our sense of the term. However, in some cases at least the animal pronoun takes the same pronoun stem in its free form as other third person pronouns. A similar situation is true for the inanimate pronoun in many Z languages.

Generic nouns have completely replaced "true" third person pronouns in some languages, such as MIX. The replacement process can be seen taking place in SI where the word for ´person´, **bɨn**, is commonly used as a third person pronoun instead of the third person pronoun.

In some Z languages generic nouns function as pronouns in certain constructions, especially relative clauses (see Sect. 5.3.1).

While the generic nouns in YTZ are also generally used as simple nouns, in TEX their distribution is much more limited. In TEX they must be the head of a relative clause, or be modified by a deictic element, or occur in a compound word. They are not phonologically independent words, but must cliticize to a following word. The generic nouns which occur in TEX are given in (25).

(25)	**Generic Noun**	**Meaning in m.s.**	**Meaning in f.s.**
	biñ	3 resp f/deity	3 resp (m/f/deity)
	yu	3 m	(not used)
	bañ	3 anim	3 anim
	fiñ	3 fam f	3 fam (m/f)

Finally, in Ch languages the generic nouns **yu** (3m), related to **ki?yu** ´man´, and **ne?** (3f), a word for ´people´ in some variants of Ch, appear in front of certain nouns, functioning almost as classifiers, it seems. (In other variants of Ch, the complementizer nu occurs in this position.) The following examples are from NC.

(26) a. ne? kuna?ã kã? ´that woman´
3f woman there

b. yu ki?yu kã? ´that man´
3m man there

c. ska yu štyũ ´a Tiltepecan (man)´
one 3m Tiltepec

3 Semantic distinctions in pronouns

Z languages all make a morphological distinction for three persons, and in third person most (but not all) distinguish among three "genders": human, animal, and inanimate (see Appendix G). Most Z languages distinguish between inclusive and exclusive in first person plural (the only exception is GL).

Most Z languages have a respect/familiar contrast in the third person. (The Ch languages do not have such an established system.) The respect pronoun may be the same as or different from the pronoun used to refer to deities. Some have respect forms in the second person singular as well (usually through the inclusion of some extra morpheme). The "respect" parameter is fairly complex in some Z languages. The use of respect forms is determined by a number of factors, including primarily the relative age of the speaker and the person referred to, the kinship relationship between them, their sex, and the attitude of the speaker towards the person. Because of the intricacy of these parameters, it may be incorrect to talk of the distinction as being between respect/formal/honorific and familiar/informal/nonhonorific.

Some Z languages display a difference between male speech and female speech for third person respect and familiar pronouns. The Ch group as well as some others have distinct forms depending on the sex of the third person referent.

Finally, in CHO an additional distinction (which is being lost, nevertheless) is made between the normal third person respect pronoun and the third person pronoun used for an elderly person. GL and IST have a pronoun to refer to a child specifically, although in IST it is obsolete. GH and OCO both have a third person pronoun with indefinite (or vague) reference.

With such a variety of third person pronouns, one might ask what happens when the intended referent is a conjunction of two groups, as "they" referring to a man and an animal. Unfortunately, I do not have the answer to this question on a cross-linguistic basis. In a text of one language (YLG) it was observed that the pronoun used in such a case was the third person respect form, which was used to refer to the man previously.

4 Number marking

All of the Z languages distinguish singular and plural in first and second person pronouns. Some Z languages also have special verb stem morphology which occurs when the verb has a first person plural subject. In other languages, the stem which occurs with a first person plural (inclusive or exclusive) subject is different from that stem which occurs with other subjects in that particular aspect. It is possible that in some cases an isolatable morpheme is involved. (In IST, and perhaps the other languages, only certain verbs have this morphology, and there is variation between speakers with respect to the verbs which use it.)

(27) a. (CHI) ri-ǰyapa-nu? (other stem: -apa)
HAB-watch-1p,incl
´we watch´

b. (IST) ri-dusi-du (other stem: -asi)
HAB-sleep-1p,excl
´we sleep´

c. (IST) ri-de?-du (other stem: -e?)
HAB-drink-1p,exc
´we drink´

The situation regarding the means for indicating plural in third person is more complex. I turn to this now. The following description is adapted from Marlett and Pickett (1985).

There are three "features" in which Z languages differ: they either have a given feature or they do not.

A. A plural morpheme (proclitic) with nominals (including pronouns).
B. A prefix on the verb (to pluralize the subject).
C. Plural clitic floating.

The languages included in the present study display all of the eight possible combinations of these three features. For example, some languages have none of these features. One has all three of them. This is shown below.

(28)

Type	Feature A	Feature B	Feature C
1	yes	no	no
2	no	yes	no
3	no	no	yes
4	yes	yes	no
5	yes	yes	yes
6	no	yes	yes
7	yes	no	yes
8	no	no	no

The majority of the languages have feature A. The Chatino languages do not, as do not the languages in the area of Miahuatlán: AMA, COA (as far as I can tell from the sources), MIA, MIX, and QUI. Several of the languages spoken in the ex-district of Villa Alta also do not have this feature, but I discuss these in more detail below.

Feature A is the use of a plural morpheme before the head noun of a noun phrase. In RIN the morpheme which now occurs with noun phrases, as opposed to bound pronouns, is etymologically the word for ´all´ and so is not cognate with the plural morpheme in other languages.

(29) a. (IST) ka yaga ´trees´
PL tree

b. (TEX) de moz ´workers´
PL worker

c. (RIN) yugu? yaga ´trees´
PL tree

d. (CHI) ra? bwi?inn ´people´
PL person

In the same way, the pluralizer precedes the (bound) pronoun which it pluralizes, wherever this pronoun is found.

(30) a. (IST) b-yaba-[ka-be]
COMPL-fall-PL-3hum
´they fell´

b. (IST) b-i-di?-du-[ka-ni] laa-[ka-be]
COMPL-ThV-give-1p,exc-PL-3inan stem-PL-3hum
´we gave them to them´

c. (RIN) ri-le?e-da? le-[gaka-nu]
CONT-see-1s stem-PL-3f
´I see them (f)´

d. (RIN) ru-yu-[gaka-nu]
CONT-watch-PL-3f
´they (f) watch´

Feature B is the use of a verbal prefix which indicates that the subject is plural. This prefix occurs only if the subject is third person. In RIN it may occur in lieu of or in addition to the plural proclitic before the subject noun phrase.

(31) a. (RIN) gu-la?-ginaX yugu? yaga
COMPL-PL-fall PL tree
´the trees fell´

b. (RIN) gu-la?-gu?-ä-ba? bɨa?
COMPL-PL-hit-3m-3anim horse
´they (m) hit the horse´

c. (RIN) tu?-yu-[gaka-nu]
CONT.PL-watch-PL-3f
´they (f) watch´
/ru+la?/ > [tu?]

d. (TAB) ǯe-le-le?e-nä?
CONT-PL-see-3resp
´they see´

e. (YTZ) ǯ-se?-ene-bo?
CONT-PL-hear-3fam
´they hear´

f. (YTZ) ǯ-EsE?E-sed geyE? bidao?
CONT-PL-study five child
´five children are studying´

Feature C involves plural clitic floating. Of course, in order to have clitic floating, a language must have a plural clitic. But most of the languages which have clitic floating have a plural morpheme which occurs only with pronouns. Therefore I classify them as **not** having feature A. In languages with clitic floating, the plural clitic does not always occur contiguous to the nominal which it is pluralizing. Rather, it always occurs in a position following the head of the constituent to which the NP belongs. That is, in a prepositional phrase such as (32), the plural follows the prepositional head. In a possessed noun phrase such as (33) and (34)--recall that free pronouns have the same structure as possessed nouns--the plural follows the nominal head. In a clause, the plural follows the verb. But it is only in the latter in which the plural morpheme does not also actually occur contiguous to the nominal which it pluralizes. A more detailed discussion of such cases follows.

(32) (YTZ) če-ga?ak-e?   ´of them´
of-PL-3resp

(33) (YTZ) žXeid-ga?ak-e?   ´their hen(s)´
hen-PL-3resp

(34) (YTZ) lle-ga?ak-e?   ´they/them (resp.)´
stem-PL-3resp

Example (35) shows a plural subject pronoun. The plural morpheme occur both contiguous to the verb and also to the pronoun.

(35) (YTZ) polisia ka? b-e?š-ak-e? to be?nne
police those COMPL-grab-PL-3resp a person
´The police arrested a person.´

The plural morpheme does not occur contiguous to a bound object pronoun, however, since it must cliticize in the post-head position. This is shown in the examples in (36). There may be multiple sources for the one plural morpheme which may occur, as the ambiguities below indicate.

(36) a. (RS)

ri-le?e-gaka-dǎ?-nu
CONT-see-PL-1s-3f
´I see them (f)´

b. (RS)

ri-le?e-gaka-dǎ?-nu
CONT-see-PL-3m-3f
´he sees them (f)´

c. (YTZ)

y-osež-ga?ak-bo?o-b
POT-untie-PL-3fam-3anim
´s/he is going to untie them´

d. (YTZ)

gw-sa?-aklen-ga?ak-e?e-bo?
COMPL-PL-help-PL-3resp-3fam
´they helped them´

e. (YTZ)

b-nežXw-ga?ak-bo?o-ne?e-b
COMPL-give-PL-3fam-3resp-3anim
´s/he gave it/them to him/her/them´

f. (YLG)

b-le?-ak-la-ba?
COMPL-see-PL-1s-3anim
´I saw them´

g. (YLG)

b-go?š-ak-a-be-n
COMPL-hand-PL-1s-3fam-3inan
´I handed it/them to him/her/them´

h. (YLG)

b-go?š-ak-e-be-n
COMPL-hand-PL-3resp-3fam-3inan
´s/he/they handed it/them to him/her/them´

A comparison of the actual plural morphemes of the various Z languages (see Appendix H) indicates some obvious groupings. The languages which have no plural morpheme which may be used in the free pronouns are all, with the exception of TAB, in the southern part of the state of Oaxaca. The languages which use some cognate of **gaka** are all located in the north and east of the state. The central and western Z languages use a morpheme cognate to **ra.** (A couple use a cognate of **be**, which is probably related.)

The few languages which use a verbal prefix are all in the area of the ex-district of Villa Alta.

There is yet one other manifestation of plural in Z languages, but for which I have little cross-linguistic data. Some languages have suppletive stem morphology (especially for existential or positional verbs) which is dependent primarily on the number of the subject. In YTZ, for example, one verb ´be´ has the stem **-nitE?E** if the subject is third person plural. This stem also may be used if the subject is second person plural. Elsewhere the stem is **-zo**.

5 Syntax

5.1 Word order

The basic word order in all Z languages is VSO (with the exception of RIN, in which bound pronouns occur in this order, but noun phrases occur in the order VOS, as I noted above). This order generally occurs whether noun phrases or pronouns are involved and whether the latter are free or bound pronouns, as the following examples from YTZ show.

(37) a. ba-z-X-ši? žna?a-bo?o-n yetžtil-En?
already-STAT-go-buy mother-3fam-the bread-the

kate? be-bam-bo?
when COMPL-awake-3fam

´His mother had gone to buy bread when he woke up.´

b. gw-če?e-šašX polsia-n? lle?
COMPL-carry-forcefully police-the 3resp
´The police carried him away forcefully.´

c. gw-dao-b le?
COMPL-eat-3anim 2s
´It bit you.´

The position of what is commonly referred to as an indirect object varies between Z languages and often within a given Z language. A very common position for indirect objects in Z languages is after the direct object, usually following a preposition (often ´of´ or a word meaning ´face´). (See Appendix I.) The following examples are from TEX, ATE, and NC, in which marking of indirect objects depends on the verb to a large extent.

(38) a. (TEX) te?es na yã rik-ã tub consef ni-r
but now 1s POT.give-1s one advice of-2s
´But now I will give some advice to you.´

b. (TEX) na nẽ?ẽ tub rišto?o lo-r
now POT.say.1s one story face-2s
'Now I will tell a story to you.'

c. (ATE) benna-bi-ã ki-a?
COMPL-give-3fam/3anim-inan of-1s
'He gave it to me.'

d. (NC) m-da-yu tñi ?ñã
COMPL-give-3m money of.1s
'He gave money to me.'

What is semantically a beneficiary in YTZ occurs preceded by the preposition 'of'.

(39) (YTZ) gw-ži?-a? to ke no?l ki-a?
COMPL-buy-1s one of woman of-1s
'I bought one for my wife.'

In YTZ and IST, however, the indirect object of 'give' is never preceded by a preposition, and may occur either before or after the direct object. The indirect object in YTZ appears to precede the direct object usually, as in the following examples:

(40) a. (YTZ) gw-dižXw-e? nada? ši pes
COMPL-pay-3resp 1s ten pesos
'He paid me ten pesos.'

b. (YTZ) b-nežXw-e? bida?o-nE? to llibr
COMPL-give-3resp child-the a book
'He gave the child a book.'

c. (YTZ) b-nežXw-e?e-bo?o-b
COMPL-give-3resp-3fam-3anim
'He gave it to him.'

But it may occur after it if the indirect object is a free form and the direct object is an inanimate bound pronoun, as in (41).

(41) (YTZ) gw-nežXw-a?a-n llebo?
POT-give-1s-3inan 3fam
'I will give it to him.'

In other Z languages (such as OCO) there is a preference to avoid indirect objects, and periphrastic constructions of the type "I give some corn, you eat it" are used.

The locative in some languages patterns similarly to the indirect object in that it can occur as a pronoun on the verb.

(42) a. (YLG) ʒ-gwe-?e-n bit
HAB-put-3resp-3inan epazote
´He put the **epazote** (herb) in it.´

b. (YLG) bi ʒ-ak sa?a-ʒo-n
NEG HAB-can POT.walk-1p,inc-3inan
´We can´t walk on it.´

There is an important constraint such that if a noun phrase or free pronoun is used in a particular position, the succeeding positions generally cannot be filled by bound pronouns. The only exceptions to this which I have seen are in ATE, MIT, and TEX, where a bound pronoun may attach to the subject NP. In ATE the only bound pronoun that can do this is third familiar **-bi**.

(43) a. (MIT) gu-dihnn Juan-re-ni
COMPL-hit -PL-3fam
´John hit them.´

b. (ATE) ra ka enne?yu-bi
PRES.say PL person-3fam
´The people said to him...´

c. (TEX) b-di wa?a-y
COMPL-hit John-3m
´John hit him.´

If the subject has been relativized and is therefore not present superficially following the verb, the object pronoun still must be free in most Z languages.

(44) (YTZ) ben? gw-zen lle-b
person COMP-hunt 3anim
´the man who hunted it´

* ben? gw-zen-Eb

In IST, however, the inanimate bound pronoun **-ni**, the only bound pronoun which may be used for verbal objects, may cliticize to the verb even if the subject is absent.

(45) (IST) b-i?iy-u bi?ku (ni) gu-da-ni la
COMPL-see-2s dog COMP COMPL-eat-3inan Q
´Did you see the dog that ate it?´

And in the following example from YTZ, which involves an infinitival form discussed in 5.3.5, the object pronoun is bound to the infinitive.

(46) (YTZ) ša?a go-ye-b
POT.go.1s INF-tend-3anim
'I am going to tend it/them.'

I mentioned in section 1.1 that in Ch direct object pronouns (as well as certain indirect objects) are often preceded by the particle ?ɪ (NC, 'of'), under certain conditions.

(47) a. (NC) na?ā-yu ?ɪ-na
COMP.see-3m of-3inan
'He saw it.'

b. (YC) k-ye? ?ɪ-?ɪ
POT-cook.2s of-3inan
'You will cook it.'

5.2 Distribution of free and bound pronouns

I consider a verb construction with a bound subject pronoun to be unmarked, and one with a free pronoun as subject to be marked. In postverbal position the free pronoun for subject is possible in some Z languages, but not common. Bound subject pronouns usually cliticize directly to the verb, although in many Z languages a limited number of morphemes (typically adverbial) may intervene.

(48) (ATE) bi-tsina?-tegaba-ka-bi
COMPL-arrive-just-PL-3fam
'They had barely arrived...'

Pronominal objects (direct or indirect) which are **not** third person cannot be represented by bound pronouns following the verb in most Z languages. Thus, schematically, there are not clauses of the following form:

(49) * V-1s-2s but V-1s 2s

'I...you'

* V-3-1s but V-3 1s

'He...me'

ATE is exceptional in that it allows second person singular objects to be bound.

(50) (ATE) r-uadi?idzɨ-nu-a-l
PRES-talk-with-1s-2s
´I talk with you.´

Third person direct objects, on the other hand, may be free or bound, with some restrictions. Bound forms as direct object are shown in examples above, such as (41); free forms (from YTZ) are shown below.

(51) a. gw-dao-b nada? ´It bit me.´
COMPL-eat-3anim 1s

* gw-dao-b-(d)a?

b. gw-dao-b le? ´It bit you.´
COMPL-eat-3anim 2s

* gw-dao-b-o?

c. b-le?i-da? llebo? ´I saw him.´
COMPL-see-1s 3fam

d. b-le?i neto? lleb ´We saw it.´
COMPL-see 1excl 3anim

e. b-ne?XbelEn? xwan-nne? llebo? bel-En?
COMPL-give John-the 3fam snake-the
´John gave him the snake.´

The restrictions on when a third person object pronoun may be free or bound vary among the Z languages. In IST the only bound object pronoun is the inanimate one. In the Ch languages, the object pronoun is never bound to the verb (since it always has the preposition ´of´ before it). Many Z languages appear to have some restrictions on when a third person (object) pronoun can be bound to the verb. In some instances the restriction appears to be purely phonological, and/or morphological, or involving a hierarchy. Three such restrictions are given below.

(52) a. A free form must be used if the subject pronoun ends in a consonant. (YTZ)

b. A free form must be used if an identical sequence of bound pronouns would otherwise result. (YTZ)

c. A free form must be used if the subject pronoun is equal to or below the object on the hierarchy 1,2-3resp-3fam-3anim-3inan. (YLG)

Free pronouns also occur in preverbal position (see section 5.4.1).

5.3 Coreference situations

In the following sections I discuss various situations in which coreference holds between two nominals, whether in a single clause or in two clauses.

5.3.1 Relative clauses. Relative clauses follow the head noun in all Z languages. There appear to be two major strategies for relative clauses in these languages. (See Appendix J.) The first involves replacement (and fronting) of the coreferent nominal in the relative clauses by a generic noun, although the generic noun may be omitted in some languages under certain conditions.

(53) **with generic noun**

a. (YTZ) gwagwE **de?e** ǯ-yažXe
firewood thing CONT-needed
´the firewood that is needed´

b. (YTZ) to bidao? **bi?i** b-selE? no?oll-en?
a child child COMPL-sent woman-the
´a child that the woman sent´

c. (YLG) to bče?ye **da** gw-a-yib-e? trast
a scrubber thing POT-REP-wash-3resp dish
´the scrubber that she will
wash dishes with´

without generic noun

d.(YTZ) to beko? dao? zo lliž bennE?
a dog small STAT.be house person

migw č-e?enE?
friend of-3resp

´a small dog that is in the house of his friend´

The second strategy simply involves the omission of the coreferent nominal in the relative clause, but in these languages a (generally) invariant form precedes the relative clause. In some languages this form is also the general complementizer, although in others it does not occur elsewhere than in relative clauses. I refer to it as the complementizer. In some languages it is optional, as in

IST.

(54) a. (IST) bi-žo?oñe ba?du (ni) bi-?ya-lu?
COMPL-run child COMP COMPL-see-2s
´The child that you saw ran away.´

b. (IST) bi-?iy-a guna?a (ni) na-ǰi?i Ale
COMPL-see-1s woman COMP STAT-love
´I saw the woman that loves Alex.´
or: ´I saw the woman that Alex loves.´

In MIX the complementizer may be preceded by the generic noun.

Headless relatives can be formed following either strategy.

(55) a. (YLG) da ǰ-ao-ba? ´what it eats´
thing HAB-eat-3inan

b. (NC) nu ti?i ´those who are poor´
COMP poor

c. (ATE) nu ra čepa ´what Chepa said´
COMP COMPL.say

In some languages an infinitive form may follow the head noun to indicate that this modifying clause indicates a typical, even professional, state.

(56) a. (YTZ) bi?i gwitX ´sportsman´
child INF.play

b. (YTZ) bennE? gwet bennE? ´murderer´
person INF.kill person

c. (YTZ) bennE? gwetE? belE? ´butcher´
person INF.sell meat

In many, perhaps all, Zapotec languages, there exist what Joseph (1978) calls "bare" relative clauses in his discussion of Modern Greek. No complementizer or generic noun occurs. These correspond to "infinitival" relatives in English. The potential aspect occurs on the verb.

(57) a. (TEX) gyit go-y ´food for him to eat´
tortilla POT.eat-3m

b. (TEX) ri?iñ gyiky ´work to do´
work POT.do

4.3.2 Reflexive clauses. There appear to be two major types of strategies in Z languages for handling reflexivity. The first type involves no special pronouns or morphology at all. Some ambiguity is tolerated in the case of third person pronouns.

(58) a. (QUI) č-an-a? na?a
HAB-see-1s 1s
'I see myself.'

b. (QUI) č-an-ša? ša?
HAB-see-3fam 3fam
'He sees himself/him.'

In some languages, such as GH, reflexive and nonreflexive readings in third person are distinguished by the omission of the overt pronoun after the pronoun stem which occurs in object position.

(59) a. (GH) ba b-wie?-me le?e-0
already COMPL-see-3resp stem-3
'He already saw him.'

b. (GH) ba b-wie?-me le?e-me
already COMPL-see-3resp 3-resp
'He already saw himself.'

Certain words are sometimes added in some Z languages, ostensibly to help disambiguate for clauses with third person referents, but also possible in other reflexive clauses.

(60) a. (OCO) gu-dini-ž (lo mizmɨ) lɨ?ɨ-ž
COMPL-hit-3fam the same stem-3fam
'He hit himself.'

b. (NC) m-hi?ĩ ?ñã (ti kwĩ?)
COMPL-hit.1s of.1s -- self
'I hit myself.'

The second type of reflexive construction involves a reflexive pronoun or reflexive verb stem. The reflexive pronounis syntactically a direct object, and the antecedent is the subject of the clause. (To the best of my knowledge, reflexive pronouns are found only in this situation, unless used emphatically or as described for Ch in section 2.2.) Superficially the reflexive and reciprocal clauses in languages such as YTZ differ from languages such as TEX and IST in that the subject pronoun in YTZ is omitted under conditions which are discussed in more detail in section

5.3.3. Examples from these languages are given below.

(61) a. (TEX) b-dĪ ub-ā
COMPL-hit.1s self-1s
'I hit myself.'

b. (IST) kadi gu-ǯi?iba-be laka-la?a-be
not POT-lift.up-3hum REFL-stem-3hum
'He will not lift himself up (boast).'

(62) a. (YTZ) bida?o-nE? b-čeč kwim-bo?
child-the COMPL-hit self-3fam
'The child hit himself.'

b. (YTZ) b-e-yot kwin be?e-na?
COMPL-REP-kill self person-that
'That person killed himself.'

c. (YTZ) b-čog kwin-a?
COMPL-cut self-1s
'I cut myself.'

d. (YTZ) ba-ǯ-EsE?E-le?i lXwežX buř bžin? ka?
already-CONT-PL-see RECIPR burro mule those
'The burros and mules already
see each other.'

GH can also reflexivize in this way, but the reflexive stem plus bound pronoun apparently cannot be permuted as a unit, as it can in YTZ.

(63) (GH) b-wie?-gahk-me
COMPL-see-REFL-3resp
'He sees himself.'

4.3.3 Subject-possessed objects. It is characteristic of some Z languages that the subject may be omitted (and typically is) if the possessor of the direct object is coreferential with the subject. Some simple examples from YTZ are given in (65).

(64) a. (YTZ) gw-dinnE yičX-a?
COMPL-comb head-1s
'I combed my hair.'

* gw-dinn-a? yičX-a?
* gw-dinn-a? yičX

b. (YTZ) golXE š̲-llapE?Eda?o-bo?o-nE?
COMPL.remove POSS-hat-3fam-the
´He took off his hat.´

The following example is a case of backward pronominalization with the subject pronoun omitted:

(65) (YTZ) ba-ǯ-lab llibr če bida?o-nE?
already-CONT-read book POSS child-the
´The child $_i$ is already reading his $_i$ book.´

(Lit., ´(He) is reading the child´s book.´)

* ba-ǯ-lab bida?onE? llibr čebo?

The following examples also show that this omission is not obligatory:

(66) a. (YTZ) bi?i skwel ka? ǯ-EsE?E-sed-bo?
child school PL COMPL-PL-study-3fam

tarea če-ga?ak-bo?
lesson of-PL-3fam

´The schoolchildren are studying their lessons.´

b. (YTZ) ǯ-yeb žoa? če-b-En?
CONT-tend corn of-3anim-the
´He (bird) watched over his corn.´

c. (YTZ) z-edE-š̲en-a? to žin dao? či-a? ngu
COMPL-come-do-1s a work small of-1s here
´I have come to do a small work of mine here.´

Languages that omit the subject pronoun in this way also apparently obligatorily omit the subject pronoun in reflexive constructions, as shown by the following examples, whereas languages that do not omit the subject pronoun under the conditions described above do not.

(67) a. (YTZ) b-čeč kwin-a?
COMPL-hit self-1s
´I hit myself.´

b. (YTZ) ǯ-EsE?E-bažE? lXwežX-ga?ak-e?
CONT-PL-hit RECIP-PL-3resp
´They are hitting each other.´

An alternative analysis of the simplest examples of these clauses in YTZ is that the direct object is incorporated into the verb, and the bound pronoun is not functioning as possessor but as subject. Indeed, this is arguably the case for a superficially similar construction in IST, as I discuss in 5.6.8.

4.3.4 Parallel verbs. Butler (1980) describes a "parallel verb construction" in YTZ which consists of "two verbs with a similar or parallel meaning, in the same aspect" (p. 152, translation mine--SM). In addition, the subjects of the two verbs must be coreferent, and the objects, if any, must also be coreferent. The subject and object pronouns are omitted from the first verb. The two verbs cannot be separated by any word. In (68b) a pronoun copy of the subject (see section 5.4.1) is expected for other reasons after the first verb, but it does not appear due to the parallel verb construction.

(68) a. (YTZ) b-isE? g-ozX-to?
COMPL-wash:head COMPL-bathe-1excl
´We washed our heads and bathed.´

b. (YTZ) ǯ-kabey gamalielEn? b-leX
POSS-horse Gamaliel COMPL-step:on

b-lib-Eb llebo?
COMPL-kick-3anim 3fam

´Gamaliel´s horse stepped on him and kicked him.´

Likewise, the object pronoun is omitted after the first verb in (68b). The following example is from QUI:

(69) (QUI) č-upćyǎ š-ne?e-ma
HAB-cry:out HAB-say-3anim
´He cried out and said,...´

This construction has not been widely reported among Z languages.

4.3.5 Adverbial clauses. In other Z languages there are various constructions involving two verbs, the first of which is often a verb of position or motion of some type, and in which the subject pronoun is omitted from one verb or the other. It is understood that the subjects of both verbs are coreferential.

(70) a. (TEX) m-biš n-as-yu
STAT-lie STAT-sleep-3m
´He is/was lying there sleeping.´

b. (MIX) r-zob wi-na zya?ad-me
HAB-sit see-1s COMPL.come-3resp
´I was sitting watching when he came.´

4.3.6 Purpose clauses. In some Z languages the subject pronoun of the verb of a purpose clause is omitted if the main verb of the clause is the verb ´come´ or ´go´. The verb in YTZ is then inflected with the infinitival prefix.

(71) a. (YTZ) zya?a go-ye žile?
STAT.come.1s INF-tend sheep
´I came to take care of sheep.´

b. (YTZ) ša?a go-ye-b
POT.go.1s INF-tend-3anim
´I am going to tend it/them.´

c. (ATE) ¢ia-bi wi-t:ia gužtiila
POT.go-3fam INF-play tomorrow
´He is going to play tomorrow.´

In ATE, and perhaps others as well, the infinitive also occurs after the auxiliary verb ´be´.

(72) (ATE) ri-yu?u-bi wi-t:ia
PRES-be-3fam INF-play
´He is playing.´

In other Z languages the subject pronoun of the motion verb is omitted, rather than the pronoun of the purpose clause. The IST example below is especially important since the verb ´come´ has a special stem which occurs with first person singular subject (**-enda**) which does not occur here. Thus the verb ´come´ lacks any indication of a first person singular subject.

(73) a. (IST) b-e?eda g-ap-a denǰu?
COMPL-come POT-tend-1s sheep
´I came to take care of sheep.´

b. (TEX) gw-a zi?i-y baňkyug
COMPL-go POT.buy-3m animal
´He went to buy animals.´

The latter example from TEX can also be said without omitting the pronoun, but the complementizer is then necessary.

(74) (TEX) gw-a-y nu zi?i-y bañkyug
COMPL-go-3m COMP POT.buy-3m animal
'He went to buy animals.'

4.3.8 Complement clauses. Subject pronouns in object complement clauses are not deleted under coreference with a nominal in the main clause in any Z language with which I am familiar.

(75) a. (YTZ) ǰ-ene?e-bo? g-ol-bo?
CONT-want-3fam POT-sing-3fam
'He wants to sing.'
or: 'He wants him to sing.'

b. (MIX) r-ladz-me tsie-me
HAB-want-3resp POT.go-3resp
'He wants to go.'
or: 'He wants him to go.'

c. (IST) r-aka-la?ǰi-be če-be
HAB-be-heart-3hum POT.go-3hum
'He wants to go.'
not: * 'He wants him to go.'

In some Z languages, however, the pronoun can be omitted from the main clause, forcing a coreferential reading.

(76) (MIX) r-laž tsie-me
HAB-want POT.go-3resp
'He wants to go.'

In IST such a construction has the opposite reading.

(77) (IST) r-aka-la?ǰi-0 če-be
HAB-be-liver-3hum POT.go-3hum
'He wants him to go.'

5.3.8 Discourse conditions. In some Z languages and also the Ch languages it is common to find situations in which a "zero" third person bound pronoun is used. This "zero" pronoun can be pluralized and can occur with the pronominal stem used with other third person bound pronouns in IST. (Thus it is not a simple case of "gapping" in IST, although it might be in other languages.) A distinction is made between the referent of an overt third person pronoun and the referent of the zero pronoun. The examples in the previous section from IST illustrate this for a single sentence example. The distinction can be maintained for long sections of a narrative in both Ch and IST, however. Some additional examples from IST are given below.

(78) a. (IST) na-be z-e?eda-Ø giži?
HAB.say-3hum FUT-come-3hum tomorrow
'He$_i$ said that he$_j$ will come tomorrow.'

b. (IST) na-be z-e?eda-be giži?
HAB.say-3hum FUT-come-3hum tomorrow
'He$_i$ said that he$_i$ will come tomorrow.'

c. (IST) na-Ø z-e?eda-Ø giži?
HAB.say-3hum FUT-come-3hum tomorrow
'He$_i$ said that he$_i$ will come tomorrow.'

5.4 Copying.

5.4.1 Fronted noun phrase. A common characteristic of Z languages is the occurrence of a bound pronominal copy of a fronted subject noun phrase (or pronoun). The details vary from language to language. In IST overt copies occur only with fronted pronouns. In other languages the copying is found sporadically in texts and the conditions are not well understood. In some languages (YTZ, for example) a copy may also occur if other types of noun phrases are fronted. I do not have extensive information on the latter cases, however. Note the following examples of subject fronting:

(79) a. (YTZ) beko?o-nE? zeča-b na?
dog-the stand-3anim there
'The dog is standing there.'

b. (YTZ) le-ga?ak-bo? ǯ-EsE?E-sed-bo?
stem-PL-3fam CONT-PL-study-3fam
'They are studying.'

c. (TEX) orze? yu-ze? gwa-y tu?uba gyedz
- 3m-there COMPL.go-3m other town
'He went to another town.'

d. (TEX) i?i r-ni?i-ñ nu kamño?oņ
here HAB-say-3inan COMP truck

p-kad gwi?i -ñ ru
COMPL-CAUS.acquire sickness 3inan 2s

'Here it says that a truck hurt you.'

e. (YTZ) yelE? ǯ-ze-taž-e? gao-b lfalf-En?
NOM CONT-?-like-? POT-eat-3anim alfalfa-the

betE-n le-b
COMPL.kill-3inan stem-3anim

'His liking to eat alfalfa killed him.'

f. (IST) la?a-be b-i?ni-be-ni
stem-3hum COMPL-do-3hum-3inan
'He did it.'

g. (IST) li?i b-i?ni-lu-ni
2s COMPL-do-2s-3inan
'You did it.'

h. (IST) Betu b-i?ni-0-ni (without overt copy)
COMPL-do-3-3inan
'Robert did it.'

In all Z languages this copy does not occur if the fronted subject is negative or interrogative.

(80) a. (YTZ) no bidao? ka? zXE-zeǯa tnez-En?
who child PL PL-STAT.stand road-the
'What children are standing in the road?'

b. (TEX) kyu zu re
who STAT.stand there
'Who is there?'

c. (IST) tu b-i?iya Ale
who COMPL-see
'Who saw Alex?'

d. (IST) una bi?ku gu-doya?a li?i
which dog COMPL-bite 2s
'Which dog bit you?'

e. (YTZ) notono no bi-dE
no:one anyone COMPL-come
'Nobody came.'

f. (TEX) se-kyu-la zu
NEG-who-EMPH STAT.stand
'Nobody is there.'

This copy also does not occur in relative clauses.

(81) a. (YTZ) ben? gw-žen le-b
person COMPL-hunt stem-3anim
´the person who hunted it´

b. (TEX) mbeky nu ri nañ ma-ze?
person COMP STAT.be inside 3anim-there

m-bañ-yu
STAT-alive-3m

´The person who was inside it was alive.´

c. (YTZ) gwagwE de?e ǰš̲azXe
firewood thing CONT.needed
´the firewood that is needed´

5.4.2 Emphatic subject. In some languages, although perhaps not generally, a free pronoun and a bound pronoun may cooccur after the verb if the subject is being emphasized in some way. (The following examples are both negative; it is not clear whether this fact is significant or not.)

(82) a. (YTZ) bito n-ak-a? nada? gual
not STAT-be-1s 1s strong
´I am not strong.´

b. (YTZ) Juan-nE? bito n-ak-e? lle? gual
John-the not STAT-be-3resp stem.3resp strong
´John, **he** isn´t strong.´

5.4.3 Modal verbs. In Z languages certain so-called modal verbs often occur without subject person markers. These verbs might be analyzed as taking a sentential subject, in which case a person marker is not expected.

(83) a. (SDA) rahk ra?a-n len tyant
HAB.able HAB.go-1s stomach store
´I can go to the store.´

b. (YTZ) ǰ-ak ǰ-om-b žin gwal
CONT-able CONT-do-3anim work hard
´It can do hard work.´

c. (YTZ) bat selao g-on-o?o-n
when POT.finish POT-do-2s-3inan
´When are you going to finish making it?´

d. (TEX) na ornu b-laž b-do-y gyit
now when COMPL-finish COMPL-eat-3m tortilla
´Now when he finished eating (tortillas)...´

e. (IST) za-nda gu-ni-be nga
FUT-able POT-do-3hum that
´He will be able to do that.´

For many Z languages, the verb ´begin´ (as well as some others in some languages) can optionally be inflected for the same person as the subject of the complement clause. I do not know if there is any language in which the erb ´begin´ **must** have a subject following it.

(84) a. (YTZ) ǯ-zolao(-bo?) g-ol-bo?
CONT-begin(-3fam) POT-sing-3fam
´He begins to sing.´

b. (IST) bi-zulu-(be) ka-yu?unda-be
COMPL-begin-(3hum) CONT-sing-3hum
´He began to sing.´

c. (TEX) b-rugyi?(-y) r-ni?i-y
COMPL-begin-3m HAB-say-3m
´He began speaking.´

d. (TEX) mbeky ze? b-rugyi?(-y) r-iky-yu ri?iň
person there COMPL-begin-3m HAB-do-3m work
´That person began working.´

5.5 Other constructions where a noun phrase is missing

5.5.1 Imperatives. All Z languages have an imperative mood in which the subject pronoun is omitted, as the following examples illustrate:

(85) a. (YTZ) g-otE?
COMPL-lie
´Lie down!´

b. (YTZ) lle?e-g-atE?
IMPER-POT-lie
´Lie down (pl.)!´

c. (YTZ) lledo-yen žin-nnE?
IMPER-go/do work-the
´Let´s do the work!´

There are also other constructions, typically potential verb forms, which are used as mild imperatives and in which the pronoun is not absent.

(86) (YTZ) bito goll-o?
NEG POT.sing-2s
'Don't sing.'

5.5.2 Locative subject. In some Z languages a pronominalized location is not marked by an overt pronoun.

(87) a. (TEX) kyit zu škye?ey
pretty STAT.stand San:Lorenzo
'San Lorenzo is pretty.'

b. (TEX) kyit zu
pretty STAT.stand
'It (San Lorenzo) is pretty.'

In IST the inanimate pronoun would be used in these cases.

5.5.3 Unspecified object. Z is similar to English in that clauses with "unspecified" objects simply involve the lack of any object form.

(88) a. (YTZ) ǯ-ao-ǯo yogwE? za
CONT-eat-1excl all zday
'We eat every day.'

b. (YTZ) zižXo ǯ-lab bida?o-nE?
fast CONT-read child-the
'The child reads fast.'

c. (YTZ) ǯ-az-a?
CONT-plant-1s
'I am planting.'

5.5.4 Meteorological verbs. Many meteorological expressions in Z languages involve a noun phrase plus verb expression, such as in (89).

(89) a. (TEX) b-et gyey 'it rained'
COMPL-descend rain

b. (TEX) d-ru?u bi 'it is windy'
CONT-leave wind

c. (IST) nu?u bi 'it is windy'
HAB.be wind

d. (IST) ka-yaba nisagie ´it is raining´
CONT-fall rain

In some languages, however, there are also some predicates which, when referring to meteorological events, do not take an overt subject pronoun.

(90) a. (TEX) J̌i rag ´it´s very cold´
very cold

b. (IST) na-nanda ´it is cold´
STAT-cold

c. (IST) na-nda? ´it is hot´
STAT-hot

There are also some verbs, of roughly this same class, which take an inanimate subject pronoun.

(91) (TEX) b-iče-ñ
COMPL-be:evening-3inan
´it was evening´

5.6 Special clausal constructions

5.6.1 Passive. Most Z languages have a passive construction in which the "agent" is not specifiable. In several languages the passive is restricted lexically to a few transitive verbs (in OCO, fifteen to twenty verbs). The lack of productivity and problems in morpheme identification in some languages have led some investigators to claim there is no passive in a given language, although the matter is debatable in my opinion. In some Z languages passive is more productive, although the expressed subject of a passive verb (CHO, for example) in some languages must be inanimate.

(92) a. (YTZ) gw-d-ao-n
COMPL-PASS-eat-3inan
´It was eaten.´

b. (CHO) ba b-d-ao-na
already COMPL-PASS-eat-3inan
´It was already eaten.´

c. (CHO) r-d-a?o bë?la?
CONT-PASS-buy meat
´Meat is bought.´

d. (TEX) b-du-gyi-y
COMPL-PASS-hit-3m
´He was appointed.´

The following examples from YTZ illustrate active vs. passive morphology. Although Butler (1980) gives the examples in (94) as an active/passive pair, it is possible that they are actually a causative/simple pair; the fortis/lenis alternation suggests that this may be the case (see Marlett and Pickett, this volume).

(93) a. (YTZ) ǰ-ot-e?e-n
CONT-grind-3resp-3inan
´She grinds it.´

b. (YTZ) ǰ-yet-En
CONT-PASS/grind-3inan
´It is ground.´

(94) a. (YTZ) ǰ-kodE?E-ga?ak-e?e-b
CONT-yoke-PL-3resp-3anim
´He yokes them.´

b. (YTZ) ǰ-EsE?E-godE?E-b
CONT-PL-PASS.yoke-3anim
´They are yoked.´

In the last pair, notice that in the active the plural agreement is a suffix, the agreement used with objects, while in the putative passive example, the plural agreement is a prefix, the agreement used with subjects.

5.6.2 Comitative object. Many Z languages have a construction in which a nominal that is semantically a comitative functions grammatically as an object. (See Appendix K.) In most of our examples there is no semantic "patient" present, but apparently this is not necessary, as the IST examples below show. The verb in these examples carries an additional suffix (which may be homophonous with some other word in the language). Compare (95a) in which the comitative nominal is in a prepositional phrase and (95b) in which it is in object position. Although in some languages it is unclear from the morphology whether it is a direct or an indirect object, there is not any direct evidence that it is anything but a direct object.

(95) a. (YTZ) ǰ-i?-a len lle-bo?
CONT-sit-1s with stem-3fam
´I am sitting with him.´

b. (YTZ) ǯ-i?i-len-a?a-bo?
CONT-sit-COM-1s-3fam
´I am sitting with him.´

The following examples further illustrate this construction:

(96) a. (YTZ) gw-dao-len-a? bida?o-ne? žgwe
COMPL-eat-COM-1s child-the female
´I ate with the girl.´

b. (SDA) bi-ně-re-ž dunuhun lěn solohiko
COMPL.go-with-PL-3 1excl stomach zoo
´They took us to the zoo.´
Lit. ´They went with us to the zoo.´

c. (TEX) na ornu ri-nu-r ñi?i yu
now when POT.arrive-with-2s 3inan house
´Now when you arrive at the house with it...´

d. (IST) ri-gite-ne-be na?a
HAB-play-with-3hum 1s
´He plays with me.´

e. (IST) ri-gite-ne-be na?a ti pelota
HAB-play-with-3hum 1s a ball
´He plays ball with me.´

As evidence of the objecthood of the comitative, consider the following example from YLG. For the subject pronoun to be omitted, the possessed NP must be an object (see section 4.3.3).

(97) (YLG) Xače?-len š̱eX š̱eo ke-n
COMPL.go.sit-COM rock river of-3inan
´He went and sat with his rocks.´

5.6.3 Instrumental object. A noun phrase which is semantically an instrument sometimes occurs superficially as an object. Evidence of objecthood is the lack of a preposition and the fact that the instrumental determines a bound pronoun. In such cases the verb is marked with the suffix -e in YTZ, here glossed INSTR. (See Appendix K.) In YLG no verbal morphology is involved.

(98) a. (YTZ) gw-deib-e-ne? laǯE?E-nE? de?e zan
COMPL-wash-INSTR-3resp cloth-the thing many

las   nis   ya?až̌til-en?
times water soap-the

´She washed the cloth with soapy water many times.´

b. (YTZ) ǯ-yib-e-da?a-n   laǯe?   kinga
CONT-wash-INSTR-1s-3inan clothes these
´I wash the clothes with it (soap).´

c. (YLG) gw-din-a-n   yid
COMPL-hit-1s-3inan cat
´I hit the cat with it.´

d. (YTZ) lle-n   g-ot-e-ǯo-n
stem-3inan   POT-die-INSTR-1incl-3inan
´We will die from it.´

5.6.4 Indirect object. In ATE and RIN some verbs have an additional object. The suffix **-ni** (low tone) occurs on the verb in ATE and on the suffix **-i?i** in RIN. The following examples are from RIN.

(99) a. (RIN) rigišXw-?a   le   r-u?una-?
CONT.explain-1s what CONT-do-1s
´I explain what I do.´

b. (RIN) rigišXw-i?i-da?   li? le   r-u?una-?
CONT.explain-IND-1s 2s   what CONT-do-1s
´I explain to you what I do.´

(100) a. (RIN) ru-lu-?a   le   r-u?un-a?
CONT-show-1s what CONT-do-1s
´I show what I do.´

b. (RIN) ru-lu-i?i-da?   li? le   r-u?un-a?
CONT-show-IND-1s 2s   what CONT-do-1s
´I show to you what I do.´

These affixes also appear to occur in some examples in which the analysis is less clear.

(101) (RIN) rinab-e?e-da?   le-gaky-ǎ?
CONT.govern-IND-1s   stem-PL-3m
´I govern them.´

5.6.5 Oblique object. A few verbs in some Z languages require the "patient" to be preceded by a preposition.

(102) a. (YTZ) leǯgwale ǯ-ake ƶ̌na? bida?o-nE? če-bo?
much CONT-feel mother child-the of-3fam
'The child's mother loves him a lot.'

b. (YTZ) b-ezi?iƶ̌en ƶ̌na?a-bo?o-n če-bo?
COMPL-forgive mother-3fam-? of-3fam
'His mother forgave him.'

c. (QUI) na?a p-tedy čeñ-ša?a
1s COMPL-hit of-fam
'I hit him.'

In OCO and MIT the stimulus which occurs in a clause with the verb 'see' is preceded by the preposition 'to' ('face'). (In OCO, if the same verb 'see' occurs with a superficial direct object (that is, without the preposition 'to'), the meaning is 'take care of'.)

5.6.6 Dative subject. There are verbs in some Z languages that occur (obligatorily or optionally) in what might be analyzed as "dative subject" constructions. The "patient" or "stimulus" is in subject position, and the "agent" or "experiencer" is in indirect object position (preceded by **ni** 'of' in TEX which marks the indirect object of the verb 'give').

(103) a. (TEX) b-yad-ñi nẽ
COMPL-find-3inan of.1s
'I found it.'

* byad-ã ñi?i
COMP.find-1s 3inan

b. (TEX) gu-?ud-ñi nẽ
COMPL-acquire-3inan of.1s
'I acquired it.'

or

gu-?ud-ã ñi?i
COMPL-acquire-1s 3inan

c. (YC) wa mso?o ?Ĩ ã
COMPL COMPL.knock:out.3 of them
'They knocked him out.'

The "agent" may be left unmentioned also.

(104) a. (TEX) b-yad-ñi
COMPL-find-3inan
'It was found.'

b. (TEX) gu-?ud-ň
COMPL-acquire-3inan
´It was acquired.´

The verbs that do this in ATE have the suffix **-ni** (with low tone--see section 5.6.4) attached to them. Among these verbs in ATE are ´see´, ´smell´, and ´hear´.

(105) (ATE) i-la?-ni ri?u ka belia
POT-see- 1incl PL star
´We will see the stars.´

The possessor of an alienable item in some Zapotec languages occurs in some non-nuclear relationship--very possibly as an indirect object--to a verb of position, as in the following examples:

(106) a. (YLG) ni?t čop beǰX
STAT.stand/PL two chicken

ke bizan-a?
of cross/sex/sibling-1s

´My sister has two chickens.´

b. (YC) ndiya ska kaha ?yā nde
???.belong one box of.1s here
´I have a box here.´

5.6.7 Locative as subject. In YLG and IST there are examples in which a locative noun occurs in subject position. The verbs in the following examples are typically intransitive verbs.

(107) a. (IST) n-anda-be bi´ga?
STAT-hang-3hum necklace
´She is wearing a necklace.´

b. (YLG) tnez-en bi ǰ-ak sa?-ǰo-n
road-the NEG HAB-can POT.walk-1p,inc-3inan

ža?a-n yeX la?ye
STAT.be:inside-3inan rock wide

´We can´t walk on the road; there are big rocks in it.´

c. (YLG) be?č-en ǰ-o?e-ba? ž̠we? bida?o
louse-the HAB-give-3anim sickness child

kate ʒ-bia-be? ba zan
when HAB-sit-3fam animal many

´Lice give a child sickness when there are many of them in him/her.´

5.6.8 Causative clauses. Most Z languages have a morphological causative construction. Basing our analysis on the languages in which there is a clear morphological difference between direct and indirect objects, we can make the following generalizations:

(108) The subject and direct object (if any) of the "embedded" verb occur as direct object(s) of the causative verb.

Examples with embedded intransitive verbs are given below:

(109) a. (YLG) b-a-s-ban-a-be?
COMPL-REP-CAUS-awaken-1s-3fam
´I woke him.´

b. (YTZ) ʒ-za-nnE?E-bo?o-b
CONT-CAUS-be:fat-3fam-3anim
´He fattened it.´

Examples with embedded transitive verbs are given below. While some syntactic theories claim that there may be only one direct object per clause, I cannot provide evidence internal to Z at this time which supports or disconfirms this claim. The ungrammatical example from TEX is included to show that indirect object morphology is not possible.

(110) a. (IST) ru-g-wa?-be-ni na?a
HAB-CAUS-carry-3hum-3inan 1s
´He made me carry it.´

(111) b. (YTZ) ʒ-go-ʒE?E-bo?o-n lle-bo?
CONT-CAUS-carry-3fam-3inan stem-3fam
´He makes him carry it.´

(112) c. (TEX) go?ow-ã-ma gyit
COMPL.CAUS.eat-1s-3anim tortilla
´I fed it (animal) tortillas.´

* go?ow-ã gyit ni-ma
of-3anim

In MIX the causative construction involving the verb ´do´ is more commonly found. The pattern is the same with respect to the arguments of the embedded clause. No indirect object morphology is found with this construction.

(113) (MIX) b-en k-lo?ob-me na? ley
COMPL-do POT-sweep-3resp 1s patio
'He made me sweep the patio.'

5.6.9 Noun incorporation. As I discussed in 5.3.3, in some Z languages the subject pronoun is omitted when it is coreferential with the possessor of the direct object. There are clauses in which this might appear to be the case in IST. These all involve body parts. Nevertheless, I follow Pickett in suggesting that these should be analyzed as instances of noun incorporation.

(114) (IST) za-gi?ibi-na ka ba?du?
FUT-wash-hand PL child
'The children will wash their hands.'

not: * za-gi?ibi na ka ba?du?
(with missing subject)

The evidence which supports the incorporation analysis is that there are adverbial affixes which may occur on the verb in IST, and these occur before the subject pronoun. In the putative cases of noun incorporation, they precede the pronoun, suggesting that the string BODY PART + PRONOUN is not a constituent.

(115) (IST) gu-di?ibi-na-ča?awi-be
COMP-wash-hand-well-3hum
'He$_i$ washed his$_i$ hands well.'

This incorporation is not obligatory, as the following examples establish.

(116) a. (IST) gu-di?ibi-ča?awi-be na-be
COMP-wash-well-3hum hand-3hum
'He$_i$ washed his$_i$ hands well.'

b. (IST) ri-ni?ib-e nye?e-sy-a?
HAB-move-1s leg-only-1s
'I simply moved my leg.'

An example of noun incorporation in Z which does not involve a possessed noun is given below.

(117) (SI) r-be?e-la-ya? bdo?
PRES-give-name-1s child
'I named the child.'

5.6.10 Impersonal intransitive clauses. In ATE there are clauses which, although intransitive semantically, involve two pronouns. One of these pronouns is obligatorily the third person nonhuman pronoun (glossed here as 3anim) and is in subject position. The other pronoun, which varies according to the reference intended, is in direct object position. The second example below has an NP in the object position.

(118) a. (ATE) gulla-ā̃ ri?u
POT.angry-3anim 1incl
'We will be angry.'

b. (ATE) r-utu?-ā̃ naga?-bi
PRES-pain-3anim ear-3fam
'His ear hurts.'

APPENDIXES

Introduction

When a given form does not exist, to my knowledge, I have written dashes in the table. If information is simply lacking, I have left a blank.

APPENDIX A: ABBREVIATIONS FOR LANGUAGES

Zapotecan (proper)

		Full name of a central **municipio**
AMA	Amatlán Zapotec	S. Cristobal Amatlán
ATE	Atepec Zapotec	S. Juan Atepec
CAJ	Cajonos Zapotec	S. Pedro Cajonos
CHI	Chichicapan Zapotec	S. Baltasar Chichicapan
CHO	Choapan Zapotec	Santiago Choapan
COA	Coatlán Zapotec	S. María Coatlán
GH	Guevea de Humbolt Z.	Guevea de Humboldt
GL	Guelavía Zapotec	S. Juan Guelavía
IST	Isthmus Zapotec	Juchitán de Zaragoza
LAC	Lachixío Zapotec	S. Vicente Lachixío
MIA	Miahuatlán Zapotec	Miahuatlán de Porfirio Díaz
MIT	Mitla Zapotec	S. Pablo Villa de Mitla
MIX	Mixtepec Zapotec	S. Juan Mixtepec
OCO	Central Ocotlán	S. Antonino Castillo Velasco
QUI	Quioquitani Zapotec	Quioquitani
RIN	Rincón Zapotec	Santiago Camotlán
SDA	Albarradas Zapotec	S. Domingo de Albarradas
SI	Santa Inés Zapotec	S. Inés Yatzeche
TAB	Tabaá Zapotec	S. Juan Tabaá
TEX	Texmelucan Zapotec	S. Lorenzo Texmelucan
WI	Western Ixtlán Zapotec	S. Ana Yareni
YLG	Yalalag Zapotec	Hidalgo Yalalag
YTZ	Yatzachi el Bajo Z.	S. Baltazar Yatzachi el Bajo
ZOO	Zoogocho Zapotec	S. Bartolomé Zoogocho

Chatino

NC	Nopala Chatino	Santos Reyes Nopala
PC	Panixtlahuaca Chatino	S. Miguel Panixtlahuaca
TC	Tataltepec Chatino	Tataltepec de Valdez
YC	Yaitepec Chatino	Santiago Yaitepec

APPENDIX B: ABBREVIATIONS FOR GLOSSES

anim	animal	INF	infinitive
CAUS	causative	INSTR	instrumental
COM	comitative	m	masculine
COMP	complementizer	m.s.	male speaker
COMPL	completive aspect	NEG	negative
CONT	continuative aspect	NOM	nominalizer
DIM	diminutive	p,PL	plural
EMPH	emphatic	PASS	passive
excl	exclusive	POSS	possessive
f	feminine	POT	potential
f.s.	female speaker	PRES	present
fam	familiar	Q	question particle
FUT	future	RECIP	reciprocal
HAB	habitual aspect	REFL	reflexive
hum	human	REP	repetitive
IMPER	imperative	resp	respect
inan	inanimate	s,SG	singular
incl	inclusive	STAT	stative aspect
IND	indirect	ThV	theme vowel

1,2,3 first, second, third (person)

APPENDIX C: STANDARDIZED ORTHOGRAPHY

ć = ts (affricate)
š = voiceless palatal fricative
ʂ̌ = voiceless retroflex fricative
X = voiceless uvular fricative
ll = fortis l (usually realized as length)
l = lenis l
nn = fortis n (usually realized as length)
n = lenis n
V?V = "broken"vowel (either glottalization or
laryngealization)
V? = checked vowel
V: = long vowel
C: = long consonant
E = shwa
ě = open e
ǎ = low front unrounded
b, d, g, ǰ, z = lenis obstruents, not necessarily voiced

APPENDIX D: FREE AND BOUND PERSONAL PRONOUNS

First person singular

	FREE	BOUND
AMA	na	-n(a)
ATE	inte?	-a?
CAJ	nede?e	-(n)a?
CHI	na?	-a?n
CHO	në?(ëdi)	-a?
COA	na, sa	-na
GH	na?a	-a?n
GL	na?a nare (object)	-a?
IST	na?a	-a?
LAC	li?-	-a
MIA	na	-n
MIT	na(ha)re	-ä
MIX	na?	-na
OCO	na	-a
QUI	na?a	-n, -a?a
RIN	ne:da?	-(d)a?
SDA	nah	-an
SI	na?a	-a?
TAB	neda?	-(d)a?
TEX	yā	-ā
WI	neeti	-a
YLG	nada?	-(l)a?
YTZ	nada?	-(d)a?
ZOO	neda?	-a?

Chatino group:

NC	na?	(nasalization)
PC	na?	(nasalization & length)
TC	na?	(nasalization & length)
YC	na?	(nasalization)

APPENDIX D, con.

Second person singular*

	FREE	BOUND		
AMA	lu	-l(u)		
ATE	lu?	-lu?	(fam)	
	kwiã?-lu?	-kwiã?-lu?	(resp)	
CAJ	le?	-o?		
CHI	yi?	-lu?	(fam)	
	yu?bi-lu?	-yu?bi-lu?	(resp)	
CHO	luë?	-(l)o?		
COA	lo:?	-lo:?	(fam)	
	ko	-ko	(resp)	
GH	li?i	-i?y		
GL	liu?	-u?	(fam)	
	la?a-bi-u?	-bi-u?	(resp)	
IST	li?i	-(l)u?		
LAC	li?i-lu	-lu		
MIA	lu?	-l(u?)		
MIT	luh	-lu	(fam)	
	gehb-lu		(resp)	
MIX	lu?	-le		
OCO	lɨ?ɨl	-l	(fam)	
	iubi-l	-iubi-l	(resp)	(one-2s)
QUI	li?i	-l, -o?o		
RIN	li?	-(n)u?	(fam)	
	kwin-u?	-kwin-u?	(resp)	
SDA	lüh	-l		
SI	lo?o	-u?		
TAB	lue?	-(n)u?		
TEX	ru	-r(u)		
WI	lu	-lu	(fam)	
	kwina-lu	-kwina-lu	(resp)	
YLG	loe	-(l)o?		
YTZ	le?	-(d)o?		
ZOO	le?	-o?		

Chatino group:

NC, PC, YC	nu?wĩ	(tone) (& length in PC)
TC	nu?ũ	------

*Some Z languages, especially Ch, use the second person plural form as a second person singular respect form.

APPENDIX D, con.

First person plural inclusive

	FREE	BOUND
AMA	be	-be
ATE	ri?u	-ri?u
CAJ	ǰo?	-ǰo?
CHI	la?-nu?	-nu?
CHO	ra?o	-ro
COA	na	-na
GH	no	-no
GL	dunnɨ	-n(ɨ)
IST	la?a-(ka)-nu	(-ka)-nu
LAC	li-?awa	-?awa
MIA	na?	-na?
MIT	nu?ur-nu	-nu
MIX	ne	-ne
OCO	lɨ?ɨ-ya?a	-a?a
QUI	no	-no
RIN	rã?u	-ru?
SDA	dunuh-re	-re
SI	la?a-nɨ?	-nɨ?
TAB	ǰi?u	-ǰu
TEX	de-ub-na	-na
WI	ri?i-tu	-ri?i
YLG	ǰio?o	-ǰo
YTZ	ǰi?o	-ǰo
ZOO		-ǰo

Chatino group:

NC	na:	-na
PC	na	-na
TC	na	-na
YC	na	-ã

APPENDIX D, con.

First person plural exclusive

	FREE	BOUND	
AMA	nu	-nu	
ATE	intu?	-tu?	
CAJ	neto?	-to?	
CHI	la?-nu	-nu	
CHO	nëto?	-ndo?	
COA	za	-za	
GH	no?	-no?	
GL	(not different from 1p inc)		
IST	la?a-du	-du	
LAC	li?i-ru	-ru	
MIA	nu?	-nu?	
MIT	nu?u-re	-nu?u-re	?
MIX	no	-no	
OCO	lɨ?ɨ-na?a	-na?a	
QUI	no?	-no?	
RIN	netu?	-tu?	
SDA	dunuhu-n	-(a)n	
SI	la?a-du?	-du?	
TAB	netu?	-ntu?	
TEX	dẽ-yã	-yã	
WI		-tu	
YLG	neto?	-to?	
YTZ	neto?	-to?	
ZOO		-to?	

Chatino group:

NC	wa-re?	-wa
PC	wa	-wa
TC	ya	-ya
YC	wa-re	-wa

APPENDIX D, con.

Second person plural

	FREE	BOUND	
AMA	gu	-gu	
ATE	lebi?i	-le	(fam)
	kwiã?-le	-kwiã?-le	(resp)
CAJ	le?	-le?e	
CHI	la?-tu	-tu	
CHO	le?e	-le	
COA	ko	-ko	
GH	la?d	-la?d, -d	
GL	la?a-t	-t(ɨ)	(fam)
	la?a-xib-tɨ	-xib-tɨ	(resp)
IST	la?a-tu	-tu	
LAC	li?i-wa	-wa	
MIA	gu?	-gu?	
MIT	la?-tu	-tu	(fam)
	gehb-tu	-gehb-tu	(resp)
MIX	to	-to	
OCO	lɨ?ɨ-tɨ(l)	-tɨ(l)	
QUI	do	-do	
RIN	libi?i-li?	-li?	
SDA	lãh-ti	-di	
SI	la?a-tu?	-tu?	
TAB	le?e	-le	
TEX	de-ru	-ru	
WI	lebi?i	-le?	(fam)
		-kwina-le?	(resp)
YLG	le?	-le	
YTZ	le?e	-le	
ZOO			

Chatino group:

NC	?ũ	-ũ
PC	(nu)?wã	-wã
TC	mã	-mã
YC	?wã	-wã

APPENDIX D, con.

Third person deity or spirit being*

	FREE	BOUND	
CHI	la?-ñi?	-ñi?	(or to show pity)
CHO	lẽ-dã?	-dã?	(spirit)
GL	la?a-ñ	-ñ	(includes sun, etc.)
SI	la?a-ʒan	-ʒan	

Chatino group:

NC	ñi	-ñi	(also respectful, nondeity)
PC	---	-ni	
TC	---	-ni	(also ancestral heroes)

---

* If the form is also commonly used for nondeity, see the next page.

APPENDIX D, con.

Third person singular (respect)

	FREE	BOUND	
AMA	me	-me	(unknown person, woman?)
ATE	le-e	-(n)e	
CAJ	ll-e?	-e?	
CHI	la?-ba	-ba	
CHO	lë-në	-në?	
COA	mä	-mä	
GH	le?e-me	-me	
GL	la?a-ž	-ž	(m.s., male referent)
	la?a-b	-b	
IST	la?a-be	-be	
LAC	li?i-nu	-nu	(god, certain people)
MIA	-------	---	
MIT	la?a-ba	---	
MIX	me	-me	(god, older adult)
	žuz	-žuz	(woman)
OCO	lɨ?ɨ-mɨ	-mɨ	(most respectful)
	lɨ?-i?	-i?	
QUI	me	-me	(conditions on usage)
RIN	lä?	-(n)ä?	(male referent)
	lä-nu	-nu	(female referent)
SDA	la?a-man	-(a)man	(m.s., primarily)
SI	la?a-ne?e	-ne?e	(also used for things)
TAB	lä?	-(n)ä?	
TEX	---	-m(i)	
WI		-bi, -e	
YLG	ll-e?e	-(l)e?	
YTZ	ll-e?	-e?	
ZOO		-e?	

Chatino group:

NC	yu	-yu	(male)
	ne?	-ne?	(female)
	hũ	-hũ	(plural)
PC	yu	-yu	(male)
		-Ø	(female)
	ne?	-ne?	(plural; sing. respect)
TC	yu	-yu	(male)
		-Ø	(female)
	ngu?	-ngu?	(plural)
YC	yu	-yu	(male)
		-Ø	(female)
	ne?	-ne?	(plural; sing. respect)

APPENDIX D, con.

Third person singular familiar

	FREE	BOUND	
AMA	ša?a	-ša?a	
ATE	la-bi	-bi	
CAJ	lle-be?e	-be?e	
CHI	la?-bi	-bi	
CHO	lë-bi?	-bi?	
COA	(same as 3 respect)		
GH	(same as 3 respect)		
GL	(same as 3 respect)		
IST	(same as 3 respect)		
LAC	li?i-ža	-ža	(moderate respect)
	li?-i	-i	(male referent)
	li?i-nžu	-nžu	(female referent)
MIA	ša?	-ša?	
MIT	la?a-ni	-ni	
	la?a-ži	-ži	(m.s.)
MIX	žo	-žo	
OCO	lɨ?ɨ-ž	-ž	
QUI	ša?	-ša?	(m.s. only, male referent)
RIN	lǎ-bi?	-bi?	
SDA	la?a-ž	-(i)ž	(m.s.)
	la?a-m	-(i)m	(f.s.) (In m.s. , it means 3fem)
SI	(same as 3 respect)		
TAB	lǎ-be?	-be?	
TEX	---	-y(u)	(m.s., male referent)
	---	-ñ(i)	(m.s., female, or inanimate referent) (f.s., male, female or inanimate referent)
WI	la-na	-na	
YLG	lle-be?	-be?	
YTZ	lle-bo?	-bo?	
ZOO		-pe?	

Chatino group: (generally the same as 3 respect)

TC	čo?	-čo?	(female)

APPENDIX D, con.

Third person singular animal

	FREE	BOUND
AMA	ma	-ma
ATE	lã	-ã (people, animals, things)
CAJ	lle-b	-b
CHI	la?-ma	-ma
CHO	lẽ-ba?	-ba?
COA	ma	-ma
GH	le?-ma?	-ma?
GL	la?a-m	-m(ɨ)
IST	la?a-me	-me
LAC	li?i?	-i?
MIA	ma?	-ma?
MIT	(same as 3 familiar)	
MIX		
OCO	lɨ?ɨ-ma	-ma
QUI	ma	-ma
RIN	lã-ba?	-ba?
SDA	la?a-b	-b
SI	la?a-ma	-ma
TAB	lã-ba?	-ba?
TEX	---	-ma
WI	(same as 3 familiar)	
YLG	lle-ba?	-ba?
YTZ	lle-b	-b
ZOO	lle-ba?	-ba?

Chatino group:

NC	?I	-?I
PC	---	-?I?
TC	ni?	-ni?
YC		-?I?

APPENDIX D, con.

Third person singular inanimate

	FREE	BOUND	
AMA	---	-y, -a	
ATE	(same as 3 animal)		
CAJ	lle-n	-n	
CHI	la?-an	-an -(i)n	(subject) (object)
CHO	le-n	-n(a)	
COA			
GH	le?e-u	-u	
GL	ɲiɲ	-ɲiɲ	
IST	la?a-ni	-ni	
LAC	li?i	-i, -e	
MIA		-i, -a	
MIT	(same as 3 familiar)		
MIX		-w, -a	
OCO	lɨ?-in	-in	
QUI		-o	
RIN	le	-0	
SDA	la?a-n	-n	
SI	la?a-nɨ	-nɨ	
TAB	lã-na	-na	
TEX	---	-ñi	
WI	(same as 3 familiar)		
YLG	lle-n	-n	
YTZ	lle-n	-n	
ZOO	lle-n	-n	

Chatino group:

NC	na	-na	
PC		-rã	
TC	na	-na	
YC	na	-na -?ĩ	 (object)

Appendix D, con.

Miscellaneous

	FREE	BOUND	
CHO	lë-gë?	-ë?	(elderly person)
COA	män	-män	(3 plural general) ?
	ša?	-ša?	(3 plural specific) ?
GH	le?e-mu	-mu	'one (indef.)'
GL	la?a-by	-by	(child referent)
IST	la?a-ni?	-ni?	(child referent)
MIX	ne		helpless person
	mgi		m.s., male of same status
	žey		man, not well known
	ni?ib		boy
	ndzo?op		girl
	ngieg		younger people (pl)
	win		child
	yi?i		sibling, contempt

APPENDIX E: REFLEXIVE AND RECIPROCAL PRONOUNS

	Reflexive	Reciprocal
AMA		
ATE	la?a + free + ba	luetsi-
CAJ	kwin + bound	lXwež-
CHI	stem + gaka + bound(3) free(1,2) + gaka	lasa:?-
CHO	kwin + bound	(none ?)
COA		
GH	gahk + bound	¢a?- (1st), sa?- (2nd, 3rd)
GL		
IST	laka + free	sa?a-
LAC	kie + bound	-----
MIA	-----	
MIT	stem + gahk (+ bound)*	
MIX	leke + free	lsa-
OCO	-----	lu-sa?-
QUI	-----	
RIN	kwin + bound	luza?a-
SDA		
SI		
TAB	kwin + bound	
TEX	ub + bound	losa?-
WI		
YLG	kwin + bound	lgwežX-
YTZ	kwin + bound	lXwežX-
ZOO		

Chatino group:

NC	-----	ta?a-
PC	ykwi? + bound	ta?a-
TC	ykwi? + bound	ta?a-
YC		ta?a-

* This reflexive pronoun typically occurs preverbally.

APPENDIX F: DEICTIC ELEMENTS IN COMPLEX PRONOUNS

	this	that	that
AMA			
ATE	COMP + i	COMP + ã?	COMP + a
	bound + i	bound + ã?	bound + a
CAJ			
CHI		ni? + ka	ni? + či?
CHO	(not used in complex pronouns)		
COA			
GH	COMP + rie?	COMP + we	COMP + kal
GL			
IST	n + ri?	n + [g]a	n + rika
LAC	ni + ngie?	COMP +	nu + kwa?
MIA			
MIT	(not used in complex pronouns)		?
MIX			
OCO	(not used in complex pronouns)		
QUI			
RIN	COMP + ni	COMP + na?	COMP + ni?
SDA			
SI			
TAB			
TEX	bound + i?i	bound + re	bound + ze?
	generic + i?i	generic + re	generic + ze?
WI			
YLG	(not used in complex pronouns)		
YTZ	(not used in complex pronouns)		
ZOO			

Chatino group:

NC	COMP + re	COMP + kwa	COMP + kã?
TC	COMP + nde	COMP + kwa	COMP + xwa, COMP + wi?
PC	COMP + nde	COMP + kwa	COMP + kã?
YC	COMP + nde	COMP + kwa	COMP + kã?

APPENDIX G: SEMANTIC DISTINCTIONS

	human/ animal/ thing	first inc/ exc	3rd resp/ fam	2nd resp/ fam	m.s./ f.s.	3m/ 3f
AMA	yes	yes		no	no	no
ATE	partially	yes	yes	nonbasic	no	no
CAJ	yes	yes	yes	no	no	no
CHI	yes	yes	yes	nonbasic	no	no
CHO	yes	yes	yes	no	no	no
COA	yes	yes	no	yes	no	no
GH	yes	yes	no	no	no	no
GL	yes	no	yes	nonbasic	yes	yes
IST	yes	yes	no	no	no	no
LAC	yes	yes	yes	nonbasic	no	no
MIA	yes	yes	no	no	no	
MIT	no	yes	yes	nonbasic	no	
MIX	yes	yes	yes	no	no	yes
OCO	yes	yes	yes	nonbasic	no	no
QUI	yes	yes	yes	no	yes	yes
RIN	yes	yes	yes	nonbasic	no	yes
SDA	yes	yes	yes	no	yes	yes
SI	yes	yes	no	no	no	no
TAB	yes	yes	yes		no	no
TEX	yes	yes	yes	no	yes	yes
WI	no	yes	yes	nonbasic	no	no
YLG	yes	yes	yes	no	no	no
YTZ	yes	yes	yes	no	no	no
ZOO	yes	yes	yes	no	no	no
Chatino group:						
(all)	yes	yes	nonbasic	(only by using plural)	no	no

APPENDIX H: NUMBER MARKING (third person)

	Before NP	In Free Pronoun	Verbal Prefix	Type (see Sect. 4)
AMA	---	---	---	8
ATE	ka	ka	---	1
CAJ	---	gak	sV?V	6
CHI	ra?	ra?	---	7
CHO	yaka	yaka	---	1
COA		---	---	8 (?)
GH	yahk	yahk	---	1
GL	de	de		1
IST	ka	ka	---	1
LAC	be?e	be?e	---	1
MIA	---	---	---	8
MIT	re	re	---	1
MIX	---	---	---	8
OCO	ra*	ra	---	1
QUI	---	---	---	8
RIN**	yugu?	gaka	la?-	4
SDA	re	re	---	1
SI	ra	ra	---	1 (?)
TAB	---	---	lé-	2
TEX	de	de	---	1
WI	ka	ka	---	1
YLG	---	ak	---	3
YTZ	---	ga?ak-	sV?V-	6
ZOO	---	ga?ak-	sV?-	6

Chatino group: type 8

* The subdialect of OCO spoken in San Antonio uses **be** instead of **ra** as the plural morpheme.

There is a closely related dialect (Rincón Sur) which uses the same morphemes but which has a type 3 system. Also, in RIN the morpheme **-pu ocurs after the relativizer if the referent is plural.

APPENDIX I: MARKING OF INDIRECT OBJECTS*

	´give´	´sell´	´say´	(other)	
AMA	order				
ATE	ki?	ki?	order	ki?	(´send´)
CAJ	order				
CHI	order		lo	lo	(´speak´)
CHO	order	kye	order	lao	(´send´)
COA					
GH	order		order	lo	(´send´)
GL					
IST	order	lu	order		
LAC	lu		lu		
MIA	lo	lo	lo		
MIT	loh/order		loh	loh/order	(´tell´)
MIX	lo	lo	lo		
OCO	order		lu		
QUI	lo	lo	lo/order	lo	(´ask´, ´tell´)
RIN	order	ke	order	order	(´pay´)
SDA	loh				
SI			lo	order	(´say´)
TAB	order	če	order		
TEX	ni	ni	lo	order	(´tell´)
WI					
YLG	order		order	ke	(´happen´, ´send´)
YTZ	order		order	ke	(´send´)
ZOO	order		order		

Chatino: ?I ´of´ in all cases except one verb of speaking whose indirect object occurs with the preposition **lo?o** ´with´.

---

* The l-initial prepositions on this page in the Z proper languages are all the word ´face´; the k- or č-initial and n-initial prepositions are all ´of´.

APPENDIX J: RELATIVE CLAUSES (with NP heads)

	W/ Generic Noun	W/ Complementizer
AMA		
ATE	---	nu
CAJ		
CHI	---	nin
CHO	optionally	---
COA		
GH	---	ni
GL		
IST	---	ni
LAC	---	nu
MIA	---	
MIT		ni
MIX	optionally + ne	ne
OCO		na
QUI	---	ni?i
RIN	---	nu (human), bi (nonhuman)
SDA	---	ni
SI	---	na
TAB	optionally	---
TEX	optionally + nu	nu
WI		
YLG	optionally	---
YTZ	optionally	---
ZOO	optionally	---
Chatino group:		
(all)	---	nu

APPENDIX K: VERBAL MORPHOLOGY

	Comitative	Instrumental
AMA		
ATE	-nu	-ni*
CAJ		
CHI	-ne	
CHO	-len	-len
COA		
GH	-no	
GL		
IST	-ne	---
LAC		
MIA	-no	
MIT	-ná	
MIX	-nie	
OCO	-nu	
QUI	-ni	
RIN	-len	
SDA	-né	-0
SI	-ni	
TAB		
TEX	-nu	
WI	-len	
YLG	-len	-0
YTZ	-len	-e
ZOO		

Chatino group:

(all)	-lyo?	---

* with high tone

References

GENERAL

Horcasitas de Barros, M. L. and Ana María Crespo. 1979. Hablantes de lengua indígena en México. Mexico City: Secretaría de Educación Pública and Instituto Nacional de Antropología e Historia.

Joseph, Brian. 1978. Morphology and universals in syntactic change: Evidence from medieval and modern Greek. Ph.D. dissertation, Harvard. Published by Indiana University Linguistics Club.

Marlett, Stephen A. and Velma B. Pickett. 1985. PluralizationinZapotec Languages. LACUS Proceedings.

Marlett, Stephen A. and Velma B. Pickett. (this volume) Syllable structure and aspect morphology in Isthmus Zapotec.

Pickett, Velma B. 1981. Unos rasgos sintácticos de las lenguas zapotecas de México. Ponencia a la Asociación Lingüística y Filológica de América Latina.

Rendón, Juan J. 1967. Relaciones internas de las lengua de la familia zapoteco-chatino. Anales de Antropología 4:187-90.

Suárez, Jorge A. 1983. The mesoamerican Indian languages. Cambridge: Cambridge University Press.

ZAPOTEC PROPER

Atepec Zapotec

Bartholomew, Doris. 1983. Gramática Zapoteca, in Nellis and Nellis 1983.

Nellis, Neil and Jane Nellis. 1983. Diccionario Zapoteco de Juárez. Mexico City: Instituto Lingüístico de Verano.

(and personal communication with Doris Bartholomew and Neil and Jane Nellis)

Cajonos Zapotec

(personal communication with Donald and Jane Nellis)

Chichicapan Zapotec

(personal communication with Joseph and Mary Benton)

Choapan Zapotec

Lyman, Larry. 1964. The verb syntagmemes of Choapan Zapotec. Linguistics 7:16-41.

Lyman, Rosemary. 1977. Participant identification in Choapan Zapotec. SIL-UND Workpapers 21: 115-32.

(and personal communication with Larry Lyman)

Coatlán Zapotec

Robinson, Dow F. 1963. Field notes on Coatlan Zapotec. Hartford Seminary Foundation.

Guelavía Zapotec

Jones, Ted and Ann Church. n.d. Personal pronouns in Guelavía Zapotec. MS.

(and personal communication with Ted Jones)

Guevea de Humboldt Zapotec

(personal communication with Donna Marks and Barbara Morse)

Isthmus Zapotec

Pickett, Velma B. 1953. Isthmus Zapotec verb analysis I. IJAL 19:292-96.

----------. 1953. Isthmus Zapotec verb analysis II. IJAL 21:217-32.

----------. 1959. The grammatical hierarchy of Isthmus Zapotec. Language dissertation No. 56, Vol. 36, No. 1, part 2.

----------. 1967. Isthmus Zapotec, in Handbook of Middle American Indians, vol. 5, ed. by N. McQuown. Austin: University of Texas Press. 291-310.

----------, et al. 1974. Vocabulario Zapoteco del Istmo. Mexico City: Instituto Lingüístico de Verano.

(and personal communication with Velma Pickett)

Lachixío Zapotec

(personal communication with Dave Persons)

Miahuatlán Zapotec

(personal communication with David and Gerry Gutierrez)

Mitla Zapotec

Briggs, Elinor. 1961. Mitla Zapotec grammar. Mexico City: Instituto Lingüístico de Verano and Centro de Investigaciones Antropológicas de México.

Stubblefield, Morris and Carol. 1969. The story of Läy and Gisaj: A Zapotec sun and moon myth. Tlalocan 6:46-62.

(and personal communication with Morris and Carol Stubblefield)

Mixtepec Zapotec

(personal communication with Roger Reeck)

Ocotlán Zapotec

Olson, Donald. 1970. The earthquake in Ocotlán. Tlalocan 6:229-239.

(and personal communication with Don and Anne Olson)

Quioquitani Zapotec

(personal communication with Michael and Laura Ward)

Rincón Zapotec

Earl, Robert. 1968. Rincon Zapotec clauses. IJAL 34:269-74.

(and personal communication with Robert Earl)

Santo Domingo Albarradas Zapotec

(personal communication with Wolfram Kreikebaum)

Santa Inés Zapotec

(personal communication with Don Olson)

Tabáa Zapotec

(personal communication with Robert Earl)

Texemelucan Zapotec

Speck, Charles. 1984. The phonology of the Texmelucan Zapotec verb. IJAL 50: 139-64.

(and personal communication with Charles Speck)

Western Ixtlán Zapotec

(personal communication with Grace Thiessen)

Yalalag Zapotec

(personal communication with Ron Newberg)

Yatzachi el Bajo Zapotec

Butler, Inez M. 1976. Reflexive constructions of Yatzachi Zapotec. IJAL 42:331-37.

----------. 1980. Gramática zapoteca. Mexico City: Instituto Lingüístico de Verano.

----------. 1982. Un relato de la hechicería en los pueblos zapotecos de la sierra en el distrito de Villa Alta. Tlalocan 9: 249-56.

(and personal communication with Inez Butler)

Zoogocho Zapotec

(personal communication with Becky Long)

CHATINO

Nopala Chatino

Marlett, Stephen A. and David Nellis. (to appear) Compendio de la gramática chatina.

(and personal communication with David Nellis)

Panixtlahuaca Chatino

(personal communication with Leslie and Kitty Pride)

Tataltepec Chatino

Pride, Kitty and Leslie Pride. 1968. Juan Cenizas in Chatino. Tlalocan 5: 358-63.

Pride, Leslie and Kitty Pride. 1970. Vocabulario Chatino de Tataltepec. Mexico: Instituto Lingüístico de Verano.

(and personal communication with Leslie and Kitty Pride)

Yaitepec Chatino

McKaughan, Howard and Barbara McKaughan. 1951. Diccionario de la lengua Chatino. Mexico: Instituto Lingüístico de Verano.

Pride, Kitty. 1965. Chatino syntax. Norman, OK: Summer Institute of Linguistics, University of Oklahoma.

(and personal communication with Leslie and Kitty Pride)

SYLLABLE STRUCTURE AND ASPECT MORPHOLOGY IN ISTHMUS ZAPOTEC*

Stephen A. Marlett and Velma B. Pickett

1 Introduction
2 Syllable structure
  2.1 Onsets
  2.2 Rhymes
  2.3 Sequence constraint
3 Aspect morphology
  3.1 Vowel-initial roots
  3.2 Regular consonant-initial roots
  3.3 Two exceptional sets of verbs
  3.4 Verbs with the theme vowel /u/
  3.5 The verb 'be sour'
4 Conclusion

1 Introduction

In this article we discuss the basic features of Isthmus Zapotec (IZ) syllable structure. Although previous work (Pickett 1967) has done this in some detail already, the present work differs from the previous analysis. First, the constituent "rhyme" is used and found useful, but "nucleus" and "syllable-final margin" are not. Second, glides are analyzed as forming part of the rhyme in some instances, whether they precede or follow the vowel. Third, glottal stop is analyzed here as a laryngeal feature of the vowel rather than as a consonant restricted to syllable-final position. Fourth, in the present analysis only glides occur after a vowel in a given syllable (except in loanwords). We then discuss the morphophonemics of aspect prefixes (and certain others) in IZ with this syllable structure in mind and present a motivated alternative to the previous nongenerative analysis (Pickett 1953, 1955). We believe this presentation also has definite advantages over previous linear generative treatments of this area in Zapotec (Speck 1984).

2 Syllable Structure

Isthmus Zapotec has the following phonemes in common native words. (We basically ignore here certain rare exceptions and also loanwords.)

Consonantal							Nonconsonantal			
Obstruents					Sonorants		Syllabic		Nonsyllabic	
p	t	č	k	fortis		nn	i	u	y	w
b	d	ǰ	g	lenis	m	n	e	o		
	s	š		fortis		ll	a			
	z	ž		lenis		l				
						r				

Vowels may occur with one of two laryngeal modifications: checked, as in **gyeʔ** 'flower'; or laryngealized, as in **zee** 'fresh corn'. We do not discuss the prosodic representation of these vowels below. It should be pointed out that laryngealized vowels occur only in stressed syllables.[1] This fact should perhaps be tied with the restriction stated below on branching rhymes (branching rhymes occur only in stressed syllables).

These sounds are grouped into syllables in very limited ways. Along the lines of Harris 1983, we propose the rules discussed in the following sections.

2.1 Onsets

Every syllable has an onset, with some word-initial exceptions. Many, but not all of these exceptions are due to the loss (historically or synchronically, in some cases) of initial /g/ (discussed below). The onset may be any of the consonants listed above, including the glides. Ignoring the exceptions, we formulate the first onset rule as follows.

Onset rule 1: Construct a tree of category O(nset)
whose obligatory branch dominates a
[-syllabic] segment.

There is a minor restriction on onsets. While fortis sonorants (phonetically long) may be onsets word-medially, they may not be word-initially: **čonnš** 'three', **belle** 'flame'.

An /s/, /š/, or nasal may be adjoined to the left of a consonantal segment. Monomorphemic words with the sibilants are rare: these are commonly found, however, on possessed forms of nouns. Superficially the /š/ typically reduces to a voiceless version of the following sonorant: /š-luuna-be/ (POSS-bed-3human) --> [Lluunabe] 'his/her bed'. If the sibilant is adjoined to a lenis (voiced) obstruent, the obstruent devoices: /š-bi?ku/ (POSS-dog) --> [špi?ku]. Since we will refer to this devoicing rule later, we formulate it here.[2]

Devoicing: [-sonorant] --> [-voiced] / [-voiced] ___

As we stated above, many words in the speech of a large number of people do not occur phonetically with an initial consonant. Dialectally these typically alternate with /g/ initial forms.

u'naa ~ gu'naa 'woman'

Nevertheless, it can be shown that the loss of /g/ is a synchronic, although not completely understood, process. The possessed form of most such vowel-initial words reveals the underlying /g/.

š-ku'naa < /š + gu'naa/

The word iza 'year' never has an initial /g/, and in fact can be shown to be a real vowel-initial word underlyingly by examining its possessed form.

š-iza < /š + iza/

The g-Deletion rule may therefore be stated as follows (although it is not completely predictable as to the forms to which it will apply): a /g/ is optionally deleted word-initially before an unstressed vowel.

g-Deletion:
(optional)

$$g \dashrightarrow \emptyset \; / \; \# \; \underline{\quad} \; \underset{[-\text{stress}]}{V}$$

A nasal may occur only contiguous to a stop (in the onset) and is homorganic with it. Other restrictions may be stated, but we are not concerned with these here. We simply state the second onset rule as follows:

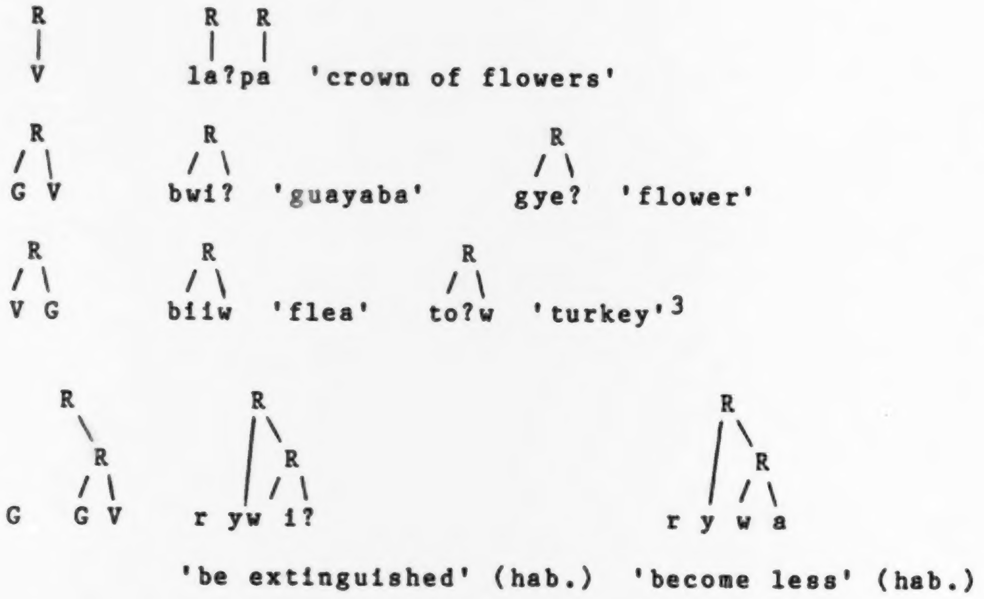
Onset rule 2: Adjoin /š/, /s/, or a nasal to the left of a consonantal segment.

Examples:

šneza	'correct'
sti	'other'
ndǐ?	'this'
[ŋ]gǎ	'that'
mbǒotǎ	'very big'
šu[ŋ]ku	'youngest child'

2.2 Rhymes

Rhymes (R) of the following types occur (V = vowel, G = glide).



R
V
R R
la?pa 'crown of flowers'
R
G V
R
bwi? 'guayaba'
R
gye? 'flower'
R
V G
R
biiw 'flea'
R
to?w 'turkey'3
R
R
G G V
R
R
r yw i?
R
R
r y w a
'be extinguished' (hab.)
'become less' (hab.)

A branching rhyme containing two glides does not occur word-initially.

These rhymes can be described by the following rules:

Rhyme rule 1:	Construct a maximally binary branching tree of category R(hyme) whose obligatory branch dominates [+syllabic, -consonantal] and whose optional branch dominates [-syllabic, -consonantal].
Rhyme rule 2:	Adjoin a [-syllabic, -consonantal] segment to the left of a rhyme beginning with a [-syllabic] segment.

Branching onsets and branching rhymes may cooccur, as the examples **stwi** 'sickness from embarrassment' and **ngiiw** 'man' show.

Counting laryngealized vowels as one, the syllable may have no more than four segments. No syllables of the type ***stwya** exist. Branching rhymes generally occur only in stressed syllables. This simple statement can be made if syllables are constructed according to the principles given above. Thus in the word **wa'ra** 'sick' the /w/ is in the onset; such a syllable may occur unstressed. In the word **kweʔ** 'next to' the /w/ is in the rhyme; such a syllable can occur only when stressed. One possible exception is the sequence /nyV/ which may occur following the stressed syllable, but not before it: **ri'zinye** 'pester' (hab.), **ru'čiinya** 'bother' (hab.). Branching onsets, on the other hand, may occur in unstressed syllables: **randa** 'be able' (hab.), **Škesa'do** 'fontanel'.[4]

2.3 Sequence constraint

Adjacent segments cannot have identical feature specification in IZ. Thus the following sequences, whether tautosyllabic or heterosyllabic, are not possible.

*iy *yi *uw *wu

Likewise there are no clusters of geminate consonants.

3 Aspect morphology

Unlike Texmelucan Zapotec (TZ) (Speck 1984), which has only four aspectual categories, IZ verbs regularly have seven--eight if the stative aspect /na-/ is counted. In certain ways the aspect morphologies of IZ and TZ are similar both with each other and also with other Zapotec languages and these similarities can be seen in the rules which are developed below for IZ. In other ways IZ is very different. Under certain conditions in TZ (Speck 1984:142) two aspect markers may occur on a single verb. A verb may also carry a negative prefix. IZ does not share these properties with TZ. More important, however, is the fact that the syllable structures of the two languages are quite different.

We discuss the allomorphy of aspect prefixes by groups of verbs. Verbs are cited in the forms they would have with a third person singular subject.

3.1 Vowel-initial roots

Of the five vowels of IZ, only two are commonly found root-initially: /a/ and /u/. Roots beginning with the other vowels are virtually nonexistent. In the following paradigms are included the only possible examples of /e/- and /o/-initial roots. The only /o/-initial root has an allomorph (which is probably the underlying form) with an initial /d/. This /d/-initial allomorph shows up in the completive aspect and elsewhere.[5] The two possible /e/-

initial roots also have stem allomorphs (see note 5). The vowel /i/ does not occur root-initially at all.

HABITUAL	POTENTIAL	FUTURE	PROGRESS.	PERFECTIVE	UNREAL	COMPLETIVE	
r-aku	g-aku	z-aku	kay-aku	way-aku	ny-aku	gu-ku	'dress oneself'
r-aana	g-aana	z-aana	kay-aana	way-aana	ny-aana	guu-na	'clean (field)'
r-a?de	g-a?de	z-a?de	kay-a?de	way-a?de	ny-a?de	gu?-de	'receive gift'
r-e?	g-e?	z-e?	kay-e?	way-e?	ny-e?	gw-e?	'drink'
r-eedž	g-eedž	z-eedž	kay-eedž	way-eedž	ny-eedž	b-eedž	'come'
r-o	g-o	z-o	kay-o	way-o	ny-o	gu-do	'eat'
r-una	g-una	z-una	kay-una	way-una	ny-una	b-ina	'obey'
r-uuna	g-uuna	z-uuna	kay-uuna	way-uuna	ny-uuna	b-iina	'cry'

We assume the following underlying forms for the aspect prefixes. Some choices of underlying forms are justified when consonant-initial roots are discussed.

r- k- z- ka- wa- ni- gb-

The following rules are needed to account for the allomorphs which occur with vowel-initial stems.

Voicing. Although it is not apparent at this point that /k-/ is the underlying form for the potential prefix, just as it is in TZ, the voiceless feature of this prefix is crucial with certain consonant-initial roots discussed below. We derive the voiced (lenis) variant by the following rule:

Voicing: k --> g / ___ + [+sonorant]

The causative prefix /k-/ undergoes the same rule.

ka + k + aǰe --> kagaǰe 'make wet' (prog.)

Gliding. A nonlow vowel becomes nonsyllabic before a vowel. Since only one syllabic position is possible in a rhyme and syllables without onsets are generally not allowed in IZ, this rule creates a segment which can be incorporated into the syllable structure. (The symbol $ signifies "syllable.")

$$\text{Gliding:}\quad \begin{bmatrix}+\text{syl}\\ -\text{low}\end{bmatrix} \longrightarrow \begin{bmatrix}-\text{syl}\\ +\text{high}\end{bmatrix} \;/\; \underline{\quad} \;[+\text{syl}]$$

Example:

```
      $                 $
     / \     |         / \
     O R     R         O   R
     | |     |         |  / \
     n i + a k u -->   n  y a k u
```

Root vowels glide before the first person bound pronoun, as in the following examples:

/r+aku+à/	-->	rakwa̋	'I dress myself' (hab.)
/r+u+too+à/	-->	rutwaa̋	'I sell' (hab.)
/r+u+dii+à/	-->	rudyaa̋	'I give' (hab.)

y-Insertion. A /y/ is inserted before other vowels which have not been incorporated into a syllable (nonincorporated vowels are indicated in a rule by a straight line above the vowel). This process is bled by Gliding, which applies to a more restrictive context.

y-Insertion:

```
                          |
     Ø --> C / ___ V
           |
           y
```

Example:

```
      $                 $ $
     /\     |           /\ /\
     OR     R           OROR
     ||     |           ||||
     ka + aku   -->     kayaku
```

Lenition. /b/ lenites to /w/ following a consonant when a nonround vowel follows, thus allowing both consonants to be incorporated into the syllable onset. (Recall that /wu/ is not a permitted sequence.)

```
Lenition:  C   C        V       ==>   C   C        V
           |   |        |             |   |        |
           g   b   [-round]           g   w   [-round]
```

Example:

```
        $               $
       / \             / \
      O   R           O   R
      |   |           |  / \
     gb + e?   -->    g  w e?
```

This rule is also used in the causative form of certain b-initial roots such as /-bezä/ 'wait'. (Causative morphology is much more irregular and unpredictable in IZ than in TZ, and we do not discuss it in detail here.) Thus /r-u-k-bezä/ becomes **rukwezä**.

Coalescence. A rhyme containing /w/ and /a/ of different morphemes reduces to /u/. The laryngeal feature of the original /a/ is retained. The sequence /wa/ does not coalesce in the prefix /wa-/ since both segments are in the same morpheme. Likewise, the sequence /wa/ internal to a root does not coalesce. This process is fed by Lenition.

```
Coalescence:     C         V          0       V
                 |         |                  |
              [-cons] +    a    ==>        [-cons]
              [+high]                      [+high]
              [+back]                      [+back]
```

Examples:

```
                              O  R            OR
                              | / \           ||
        gb + aku      ->      gw + aku   -->  gu ku

        gb + aana     ->      gw + aana   --> guuna
```

The derivation of the form **gudo** 'eat' (compl.) is unexplained at this point. We return to it later. The /d/ is an irregularity that is associated with this root in the completive (an old feature found throughout Zapotecan) (see note 4).

Dissimilation. A morpheme-initial /u/ changes to /i/ following a /b/. Then other changes occur which are discussed immediately below.

```
Dissimilation:     C     +    V             C     +    V
                   |          |      ==>    |          |
                [+lab]     [+lab]        [+lab]  [-lab ]
                                                 [-back]
```

```
Examples:  gb + una    --> gb + ina   (--> b + ina)

           gb + uuna   --> gb + iina  (--> b + iina)
```

The Erasure Convention. Harris 1983 motivates the following convention in Spanish:

> Erasure Convention: Segments not incorporated into syllable structure at the end of a derivationare erased.

We use this convention to account for the loss of the /g/ of the prefix /gb/ in those cases where /b/ does not lenite. (We also use it to describe other allomorphy below.)

Examples:

```
O    R             OR         OR
|    |             ||         ||
gb + una   -->  gbina  -->  bina

gb + eedã -->  beedã
```

The root /-eedã/ must be marked as [-Lenition], preventing the surface form ***gweedã** and forcing the loss of the /g/.

3.2 Regular consonant-initial roots

A wide variety of consonants are found at the beginning of roots. In this section we discuss what is in our analysis the "regular" set of consonant-initial verbs.

r-yaka	g-yaka	z-yaka	ka-yaka	wa-yaka	n-yaka	b-yaka	'enjoy'
r-yuu	g-yuu	z-yuu	ka-yuu	wa-yuu	n-yuu	b-yuu	'enter'
ri-wiini	gi-wiini	za-wiini	ka-wiini	wa-wiini	ni-wiini	bi-wiini	'become little'
ri-kabĩ	gi-kabĩ	za-kabĩ	ka-kabĩ	wa-kabĩ	ni-kabĩ	bi-kabĩ	'answer'
ri-dubi	gi-dubi	za-dubi	ka-dubi	wa-dubi	ni-dubi	bi-dubi	'be worn out'
ri-lã	gi-lã	za-lã	ka-lã	wa-lã	ni-lã	bi-lã	'push'
ri-niti	gi-niti	za-niti	ka-niti	wa-niti	ni-niti	bi-niti	'be lost'
ri-ree	gi-ree	za-ree	ka-ree	wa-ree	ni-ree	bi-ree	'leave'

Stems beginning with /y/ act a bit differently from other roots and so we treat them first here. In addition to the /y/-initial roots illustrated above, there are forms such as /-y-uutu/ which we take as the passive form of /-uutu/ 'grind'. These verbs, whether analyzed as (lexically restricted) passives or not, act just like /y/ initial roots.[6]

i-Deletion. As we noted in Sect. 2.3 the sequence /iy/ is not permissible. When the prefix /ni-/ occurs before /y/ roots, the /i/ deletes.

```
i-Deletion:  V C         Ø C
             | |   ==>     |
             i y           y
```

Example: ni + yaka --> n + yaka

This rule also deletes the /i/ of the causative prefix /si-/.

ka + si + ya --> kasya 'make dissolve' (prog.)

The potential prefix /k-/ voices by the rule given in Sect. 3.1.

k + yaka --> gyaka

The allomorph of the completive with these roots is accounted for by the Erasure Convention.

```
 O     R         O  R
 |    / \        | / \
gb + y aka  -->  b y aka
```

For other consonant-initial roots, we propose two insertion rules to account for the /a/ after /z-/ and the /i/ after the prefixes /r-/, /k-/, and /gb-/.

a-Insertion. An /a/ is inserted after a /z/ which has not been incorporated into the syllable structure.

```
a-Insertion:                  |
                Ø --> V /     C ___
                      |       |
                      a       z
Example:
                   OR         OROR
            |      ||         ||||
            z  +  lã    -->   zalã
```

i-Insertion. An /i/ is inserted after other consonant-final aspectual morphemes. The /i/ is *not* inserted after the causative prefix /k/. (Thus i-Insertion does not apply to the underlying form /r-u-k-dindẽ/ 'make fight' (hab.). The /k/ cannot be incorporated and deletes by the Erasure Convention after causing the following consonant to devoice: **rutindẽ**.) This rule feeds Voicing.

```
i-Insertion:                       |
                  ∅ --> V /  # (C) C ___
                        |
                        i
```

Examples:

```
             OR      OROR
        |    ||      ||||
        r + lä  -->  rilä

        k + kabï --> kikabï -->  gikabï

        gb + lä --> gbilä --> bilä
```

In Sect. 2 we mentioned a process which optionally deletes word-initial /g/ in unstressed syllables. This process applies commonly to the /g/ of the potential and completive morphemes. Crucially, it applies <u>after</u> i-Insertion. Therefore the potential form **gikabï** commonly alternates with **ikabï**.

There is one /y/-initial root which is unlike the /y/-initial roots discussed above in that it behaves in every way as if it began with some other consonant. It therefore also allows /iy/ sequences.

3.3 Two exceptional sets of verbs

Besides the "regular" set of consonant-initial verbs, there are the following two sets which are exceptional with regard to i-Insertion:

Set I:

ri-na	na	za-na	ka-na	wa-na	ni-na	gu-na	'agree to'
ri-žiji	žiji	za-žiji	ka-žiji	wa-žiji	ni-žiji	gu-žiji	'sound'
ri-žana	žana	za-žana	ka-žana	wa-žana	ni-žana	gu-žana	'give birth'
ri-dindä	tindä	za-dindä	ka-dindä	wa-dindä	ni-dindä	gu-dindä	'fight'
ri-giže	kiže	za-giže	ka-giže	wa-giže	ni-giže	gu-diže	'put'
ri-guu	ku?	za-guu	ka-guu	wa-guu	ni-guu	gu-luu	'put in'
ri-giru	kiru	za-giru	ka-giru	wa-giru	ni-giru	gu-jiru	'pinch'
ri-biji	k-wiji	za-biji	ka-biji	wa-biji	ni-biji	gu-liji	'call'
ri-bi	k-wi	za-bi	ka-bi	wa-bi	ni-bi	gu-ri	'sit'

Set II:

ri-čessa	**gi-čessa**	**za-čessa**	**ka-čessa**	**wa-čessa**	**ni-čessa**	**gu-čessa**	'jump'
ri-nã	**gi-nã**	**za-nã**	**ka-nã**	**wa-nã**	**ni-nã**	**gu-nã**	'have sight'
ri-làa	**gi-làa**	**za-làa**	**ka-làa**	**wa-làa**	**ni-làa**	**gu-làa**	'be broken'

Both of the sets of verbs illustrated above have the allomorph [gu] in the completive instead of the allomorph [b] of the previous set. We do not see presently any way to predict this fact and so we claim that the verbs of Set I and Set II are marked lexically as being exceptional to the i-Insertion rule when the completive prefix is present. We propose the following statement for Set II verbs. (Set I verbs are part of a larger pattern and are discussed below.)

> For Set II verbs: mark the form [-i-Insertion] in the context of COMPLETIVE.

The i-Insertion rule failing, we obtain representations such as the following:

```
          O R
| |       | |
g b  +  n a
```

Another rule applies which alters the prefixal sequence so that it can be syllabified (and not be lost through the Erasure Convention). The /b/ is changed to /u/.

```
Vocalization:   |           |
                C   -->     V
                |           |
                b           u
```

Thus /gb-na/ (COMPL + 'have sight') does not become ***binã**, as it would if it were regular and i-Insertion applied. Nor does it become **nã**, which it would if the Erasure Convention were to apply immediately. But it becomes **gunã** by means of Vocalization. This same situation occurs when the completive morpheme occurs before the irregular allomorph /-do/ of the root /-o/ 'eat': /gb-do/ --> **gudo** (see note 4).

Set I verbs are irregular in a more general way. The i-Insertion rule is also not applied following the potential morpheme. Since the potential morpheme and completive morpheme both end in stops underlyingly, we can state the following generalization:

> For Set I verbs, mark the form [-i-Insertion] when the root is immediately preceded by a [-continuant] segment.

Thus the potential form /g+na/ 'agree to' does not become ***gina** by i-Insertion but the /g/ is lost by the Erasure Convention, leaving **na** as the potential form. With /b/ initial stems of this class, Lenition applies after Voicing attempts to apply (and does not because these are obstruent initial roots). The application of Lenition allows the velar stop of the potential morpheme to be incorporated into the syllable and not be lost by the Erasure Convention.

```
               O  R
               | / \
k + bi -->     k  w  i
```

After Lenition applies, Devoicing (see Sect. 2.1) applies, devoicing the initial obstruent of other roots of this class. (In regular verbs, Devoicing is bled by i-Insertion.) The Erasure Convention is responsible for the ultimate loss of the potential /k/ which has not been incorporated.

Examples:

k + giru --> k + kiru --> kiru

k + dindë --> k + tindë --> tindë

Causative verbs which use the /k-/ allomorph of the causative prefix also show the same alternation. (The /u/ prefix in the following example is discussed below.)

r + u + k + dindë --> r + u + k + tindë -->

rutindë 'make fight' (hab.)

As far as the form of /b/ and /g/ initial roots in the completive is concerned, these irregular stem alternations are an old and widely attested feature of Zapotecan languages. Speck 1984 derives the initial /g/ and /b/ by special rules from underlying forms beginning with the consonants which appear in the completive forms. We suggest that similar rules are needed here although we do not formulate them.

The laryngealized vowel of /-guu/ changes to a checked vowel in the potential. Such alternations are attested elsewhere in IZ, but we have nothing to say about them here.

3.4 Verbs with the theme vowel /u/

A large number of basic roots occur lexically with the theme vowel /u/ before them.[7] Some examples are given below. At this time we have no means of predicting the set of roots which occur in this class. The majority of derived causative verbs also occur with this theme vowel, which we suggest is therefore inserted as part of the regular causative formation rule.

r-u-kaabǐ	g-u-kaabǐ	z-u-kaabǐ	k-u-kaabǐ	wa-kaabǐ	n-u-kaabǐ	b-i-kaabǐ	'commission'
r-u-dii	g-u-dii	z-u-dii	k-u-dii	wa-dii	n-u-dii	b-i-dii	'give'
r-u-yǎa	g-u-yǎa	z-u-yǎa	k-u-yǎa	wa-yǎa	n-u-yǎa	b-i-yǎa	'dance'
r-u-zelò	g-u-zelò	z-u-zelò	k-u-zelò	wa-zelò	n-u-zelò	b-i-zelò	'belch'

Causatives:

r-u-k-wezǎ	g-u-k-wezǎ	z-u-k-wezǎ	k-u-k-wezǎ	wa-k-wezǎ	n-u-k-wezǎ	b-i-k-wezǎ	'make wait'
r-u-si-dubi	g-u-si-dubi	z-u-si-dubi	k-u-si-dubi	wa-si-dubi	n-u-si-dubi	b-i-si-dubi	'make worn out'

We propose that verbs such as /-u-dii/ occur in the lexicon with the /u/ as a prefix rather than as part of the root partly because of the parallelism with the /u/ which occurs in obviously derived verbs (the causatives) in which the /u/ is not part of the root, and also so that we may maintain the generalization regarding stress assignment. Stress is assigned to the first vowel of the root. If the /u/ of these verbs were to be considered as a root-initial vowel, this generalization would have a large number of exceptions instead of the two or three which now exist. Finally, the /u/ is clearly a separable morpheme since in the causative forms of such verbs the /u/ does not occur before the root: **r-u-zunisa** 'sweat', **r-u-sunisa** 'make sweat'.

This /u/ changes to /i/ by the same dissimilation rule which affects root-initial /u/.

Example: gb + u + kaabǐ --> gbikaabǐ --> bikaabǐ

This pretonic vowel /u/ creates problems for the three aspect prefixes which have vowels since they cannot be syllabified straightforwardly. Recall that branching rhymes are generally allowed only in stressed syllables and that every syllable generally must have an onset.

```
OR        OR       OR        OR       OR        OR
||    |   ||       ||    |   ||       ||    |   ||
ka + u + dii       wa + u + dii       ni + u + dii
```

We first discuss the unreal prefix /ni-/. Gliding applies and the following intermediate structure is obtained:

```
O R
|/ \
ny u dii
```

But branching rhymes are generally not allowed in unstressed syllables in IZ. The /y/ of the unreal prefix simply deletes.

```
y-Deletion:     R  -->  R
               / \      |  where R occurs in an
               y V      V  unstressed syllable
```

Thus underlying /ni-u-dii/ becomes **nudii**.

It is not clear at this point what processes are involved in the derivation of the other forms with vowel-final prefixes. For some reason, y-Insertion does not apply.

ka + u + ==> ku

wa + u + ==> wa

It should be noted that if the /a/ were to delete after /w/, as it does after /k/, the non-permitted sequence /wu/ would be generated.

3.5 The verb 'be sour'

The following verb appears to be exceptional in several ways.[8] Word stress is pronounced on the checked vowel.

ri-ī?   gi-ī?   sa-ī?   ka-ī?   wa-ī?   ni-ī?   bi-ī?   'be sour'

The allomorphy which occurs with the verb 'be sour' is exactly that which occurs with regular consonant-initial roots, although it does not begin with a consonant. The epenthetic vowels /i/ and /a/ occur after the appropriate prefixes. There is no gliding of the /i/ of /ni-/. There is no epenthetic /y/ after /ka-/ and /wa-/. These forms also appear to be exceptional in that they have syllables without onsets. If this were an /i/-initial root, it would be the only one attested in the language. Finally, the causative form is (cited in the habitual aspect) **rusiī?**. The allomorph /si-/ of the causative commonly occurs with consonant-initial roots.

We propose an analysis here in which these exceptional characteristics are accounted for by one device: the verb 'be sour' has an empty root-initial consonant position. (A similar analysis has been proposed for larger classes of verbs in other languages, such as Seri (Marlett and Stemberger 1983).) The proposed lexical form of this verb is as follows:

```
C V    'be sour'
  |
  i?
```

The following derivations show how this representation accounts for the surface forms.

```
C    CV          CV CV
|     |          ||  |
r +  i?    -->   ri  i?

CC   CV          CCV CV         CV CV
||    |          |||  |         ||  |
gb + i?    -->   gbi  i?  -->   bi  i?

C     CV         CV   CV
|      |         ||    |
z  +  i?   -->   za    i?

CV    CV
||     |
ka    i?
```

Such an example provides additional evidence for the claim that the CV skeleton and the phonemes of a morpheme need to be represented on two separate tiers.

4 Conclusion

In this brief work we have shown that a number of phonological processes in IZ interact with the extremely restricted syllable structure that this language has (compared, for example, with TZ). Furthermore, we have motivated underlying forms for even prefixes such as the completive morpheme. (This is interesting to the first author personally since the first course he took in generative phonology used the IZ /gu-/ ~ /bi-/ alternation as an example of clear suppletive allomorphy.) However, not all exceptional classes have been eliminated, and the presence of a large class of verbs taking the theme vowel /u/ is unexplained. Some, but not all, Zapotec languages have a class of verbs with the theme vowel. TZ does not.

We hope that further investigation of other Zapotec languages may shed light on this curious fact.

Notes

*We thank our colleague María Villalobos for confirming the Zapotec data which were not drawn from previous publications. The contribution made by the first author was facilitated by a research position at the Universidad Nacional Autónoma de México this past year.

1. We represent words below in their utterance-medial form. Laryngeal modification of unstressed vowels occurs only prepausally and is predictable from other information. Pickett (et al.) 1965 gives the prepause form. In this regard, see also Mock 1982. Stress generally occurs on the first vowel of the root. Exceptions (nouns, primarily) are indicated in our transcriptions by means of an apostrophe preceding the stressed syllable. We indicate the superficially contrasting tones in this article in the following way: high tone [á], low rising tone [ǎ], and low tone (unmarked).

2. When this possessive prefix is added before a root beginning with a strident segment, the following changes occur: š + š, ž, č, ǰ --> š, š + z --> s. (No examples of š + s have been found.) These changes can be stated formally as follows:

$$\begin{array}{ccccc} [+\text{strid}] & \begin{bmatrix}+\text{strid}\\ \alpha\text{high}\end{bmatrix} & & [\alpha\text{high}] & \\ 1 & 2 & \Longrightarrow & 1 & \emptyset \end{array}$$

3. Phonetically a rhyme such as /o?w/ is pronounced as [ow?].

4. If /ny/ were analyzed as a simple palatal consonant, this problem would disappear. The reason such a move is questionable is that Cy clusters are common and /ny/ fits into this pattern straightforwardly.

5. A few verbs have a different stem when the subject is first person plural (exclusive or inclusive). For example, the verb 'sleep' is /-dusi/ with such a subject, but it is /-asi/ generally (and for some speakers even with a first person plural subject). A couple of verbs have different stems when the subject is first person singular. The verb 'eat' is /-do?/ when the subject is first person plural (in all aspects), /-dawa?/ when the subject is first person singular in the completive aspect, /-awa?/ when first person singular in other aspects, /-do/ for other subjects in the completive aspect, and /-o/ elsewhere. (We discuss the special completive aspect stems below.) The two /e/-initial

verbs listed here both have other stems: 'drink' is /-de?/ with first person plural subject, and /-e?/ otherwise; 'come' is /-yupa/ with first person plural subject, /-enda/ with first person singular subject, and /-eeda/ otherwise.

6. Previous analyses of IZ have claimed implicitly that there are no passive forms, although (usually heavily restricted) passives have been described in other Zapotec languages (Speck 1984). We suggest here that there are at least several passive forms in IZ. The following examples, which were previously analyzed as simple/causative verb pairs, illustrate the morphology involved with passives:

	Active	Passive
'grind'	r-uutu	r-y-uutu
'grab'	r-aša	r-y-aaša
'rub'	ri-gaabi	ri-daabi
'slice'	ri-guza	ri-luuzš
'shake'	ri-gibi	ri-biibi

The consonant-initial roots above are all verbs which have a special stem in the completive of the active (see Sect. 3.3); this completive stem generally begins with the same consonant as that which appears in the passive stem.

7. The idea of a theme vowel is suggested in unpublished work by W. Kreikebaum and D. Tuggy (Kreikebaum 1983).

8. Another verb which is irregular lexically in that it has a vowel cluster underlying is /-ui? diiǰa/ 'chat'. This verb conjugates in every way like a /u/-initial verb although the /u/ is not syllabic: [rwi?] (hab.). Stress falls on the /i/. The completive stem is /-yui?/. We list the surface forms of this verb below.

rwi? gwi? zwi? kaywi? waywi? nywi? bywi?

References

Harris,James W. 1983. Syllable structure and stress in Spanish: A nonlinear analysis. Cambridge: MIT Press.

Kreikebaum, Wolfram. 1983. Morphological causatives in Zapotec of Santo Domingo Albarradas. Manuscript.

Marlett, Stephen A. and Joseph Paul Stemberger. 1983. Empty consonants in Seri. Linguistic Inquiry 14:617-639.

Mock,Carol C. 1982. A systemic phonology of Isthmus Zapotec prosodies.Paper presented at the Ninth International Systemic Workshop, York University, Toronto.

Pickett,Velma B. 1953. Isthmus Zapotec verb analysis I. IJAL 19:292-296.

Pickett,Velma B. 1955. Isthmus Zapotec verb analysis II. IJAL 21:217-232.

Pickett, VelmaB. 1967. Isthmus Zapotec. Handbook of Middle American Indians, vol. 5, ed. by Norman A. McQuown, 291-310. Austin: University of Texas Press.

Pickett,Velma B., et al. 1965. Vocabulario Zapoteco del Istmo, 2 ed., rev. Mexico City: Instituto Lingüístico de Verano.

Speck, Charles H. 1984. The phonology of the Texmelucan Zapotec verb. IJAL 50:139-164.

ÑUMÍ MIXTEC SYLLABLE STRUCTURE AND MORPHOLOGY*

Laura Gittlen and Stephen A. Marlett

1 Introduction
2 Word structure
3 Phonemes
  3.1 Consonants
  3.2 Vowels
4 Syllable structure
  4.1 Onsets
  4.2 Rhymes
5 Verb morphology
  5.1 Nonfuture tense prefix
  5.2 Future tense prefix
  5.3 Other prefixes
6 Compound words
7 Motion verbs
8 Clitics

1 Introduction

Mixtec syllable structure as well as certain aspects of the phonology have been described in a variety of works for various of the Mixtec languages (Alexander 1980, Bradley 1965, Hunter and Pike 1969, North and Shields 1977, Pankratz and Pike 1967, Pike and Cowan 1967, Pike and Ibach 1978, Zylstra 1980). Our treatment of Ñumí Mixtec (NM) syllable structure is brief but crucial to what follows. In the body of this paper we discuss the morphology and segmental morphophonemics of Mixtec in a degree of detail which has never been done before. The most novel aspect of our work below, for Mixtec studies, is that we posit underlying forms for the two tense prefixes and then account for the variety of surface forms. Since in many--indeed, in a majority--of situations the underlying segments are not present in the surface form, the analysis is not immediately obvious.

2 Word structure

A minimal independent word in NM typically consists of two syllables. This seems to be a common feature of Mixtec languages. (See the works cited above.) This is true for both nouns and verbs, as the following examples illustrate.

(1) a. kiti 'animal'
b. kata 'to sing'[1]

There are a number of clitics, all of which are monosyllabic, and some adverbs and sentence level functors which have only one syllable. Compounds and loanwords account for all nouns of more than two syllables. Compounding and affixation, discussed below, account for verb forms which are more than two syllables in length.

Verb forms are always morphologically complex in that a verb is always inflected for either the future tense or nonfuture tense[2] (discussed below), although the tense prefixes do not always appear overtly (except by tone sandhi). The tense prefixes appear overtly in the following form:

(2) a. xi + to 'to take care of', NFUT
b. k^w + iso 'to carry', FUT

Clitics may precede and follow these forms, as in the following example:

(3)	[ni]	[ka]	[xi + či]	[ni]
	ASPECT	NUMBER	TENSE + ROOT	PERSON
	completive	plural	NFUT + 'bathe'	2 pl
	'you (pl.)	bathed'		

The verb word itself may be lengthened by the addition of the repetitive morpheme /^{n}da-/. According to the analysis which we present below, verbs with this morpheme always have the following structure, although there will be no overt segmental realization of the tense morpheme at the beginning of the word. We give the proposed underlying form (UF) as well as the surface form (SF).

(4) a. $\#\#\,[_V\ \text{TENSE + REPETITIVE}\ \ \#\,[_V\ \text{FUTURE + ROOT}\,]_V\ \#\ \ ]_V\ \#$

b. UF: ##k^w-^{n}da#k^w-ate##
SF: ^{n}dakate 'to wash'

c. UF: ##xi-^{n}da#k^w-či?i##
SF: ^{n}dači?i 'to replant'

Primary word stress always falls on the penultimate syllable of the word in NM. In the following formalization of this generalization, X^0 is understood to vary over all word-level syntactic category labels.

(5) Stress Assignment: V ---> [+ stress] / ___$(C_0)V]_X$0

One phonetic process refers crucially in its environment to the word boundary. A glottal stop is inserted before a word-initial vowel, even when the word occurs in the middle of an utterance. This is shown in (7).

(6) Glottal Insertion: Ø --> ? / # ___ V

(7) [?]a [?]oo ^{n}deyu [?asĩĩ]
Q NFUT-to=be food tasty
'Is there tasty food?'

The glottal stop is inserted in compound words such as **ti#[?]ina** 'dog' and **^{n}di#[?]iši** 'spider'. The structure of these words is discussed below.

3 Phonemes

3.1 Consonants

The following consonant phonemes occur in native words in NM, although those in parentheses occur only rarely and we do not discuss them further.[3]

(8)

		t		k	k^w
		^{n}d			
	v	(d)	č	x	
		s	š		
	m	n			
		l			
		(ř)			
	w		y		

These symbols have their normal phonetic values except that /v/ represents a voiced bilabial fricative and /d/ is a voiced interdental fricative.

Two phonemes have notable allophones (although not atypical for Mixtec). /t/ is released nasally before a nasalized vowel or glide.

(9) a. /tuni/ [t^{n}ũnĩ] 'mole, spot'
b. /tu?ũ/ [t^{n}ũ?ũ] 'word'
c. /tyuũ/ [t^{n}ỹũũ] 'task'

The phoneme /y/ is always realized as a nasalized glide before a nasalized vowel as in (9c) above and in **iyũ** [ĩỹũ] 'six'. Otherwise, syllable-initially /y/ is pronounced as a voiced alveopalatal fricative: **yu?u** [žu?u] 'to be afraid'.

In the remaining environment, which is following a consonant and preceding an oral vowel, /y/ is [y]: **syu?u** [syu?u] 'to frighten'.[4]

The phoneme /^{n}d/ exhibits free fluctuation between [^{n}d] and [t] in the environment following /š/, as in **š^{n}da?u** [š^{n}da?u] ~ [šta?u] 'to cheat'.

3.2 Vowels

The vowel phonemes are /i, e, a, o, u/. These vowels occur contrastively both checked (V?) and nasalized. A vowel may also be nasalized and checked. Checked vowels occur only in the stressed syllable.

(10)						
a.	^{n}de?.yu	'mud'	d.	i.ni	'inside'
b.	^{n}de.yu	'food'	e.	yu?.ma	'smoke'
c.	i?.ni	'heat'	f.	yu.ma	'wax'

Nasalized vowels, with very few exceptions,[5] occur underlyingly only in the final syllable of the word. Within the word, vowels and glides before a nasalized vowel or glide or a nasal consonant are nasalized phonetically. Nasalization spreads regressively within the word until stopped by an obstruent. A vowel **following** a nasal consonant is also slightly nasalized. We differentiate in the following chart between glides (G), nasals (N), and other consonants (C):

(11)				
	CV?Ṽ	ka?ã	[kã?ã]	'to say'
	CVCṼ	saxĩ	[saxĩ]	'nephew'
	VGṼ	iyũ	[ĩỹũ]	'six'
	CVṼ	kuũ	[kũũ]	'four'
	CGVṼ	tyuũ	[t^{n}ỹũũ]	'task'
	CGV?Ṽ	čyuũ	[čỹũũ]	'chicken'
	GVGṼ	yuyũ	[ỹũỹũ]	'bee colony'
	GVṼ	yu?ũ	[ỹũ?ũ]	'land'
	GVNV	yunu	[ỹũnũ]	'string bag'
	NVV	nuu	[nũ]	'face'
	NVGV	niyi	[nĩỹĩ]	'blood'
	NV?V	mi?i	[mĩ?ĩ]	'garbage'
	NV?GV	ma?yu	[mã?ỹũ]	'in the middle of'
	NVNV	nama	[nãmã]	'soap'
	GV?NV	ya?mi	[ỹã?mĩ]	'sweet potato'
	CV?NV	ta?nu	[t^{n}ã?nũ]	'bent'
	CVNV	kana	[kãnã]	'to shout'
	NV?V	ni?i	[nĩ?ĩ]	'to find'
	NVCV	(not found)		

Nasal spreading does not cross word boundaries as the following examples involving compounding illustrate:

(12) a. ke#ne?e [ke#nẽ?ẽ] 'get, obtain'
b. sa#yiĩ [sa#ỹĩĩ] 'corncob'
c. yũ#yivi [ỹũ#živi] 'everyone'
d. tũ#yuxa [t^{n}ũ#žuxa] 'pine tree'

The phoneme /^{n}d/ does not condition allophones, as the nasal phonemes /m/ and /n/ do. This is one of the reasons for treating /^{n}d/ as a non-nasal phoneme. Contrast the following examples, where /^{n}d/ acts as a non-nasal[6]:

(13) a. xa?^{n}de [xa?^{n}de] 'to cut'
b. xa?nu [xã?nũ] 'to grow'

(14) a. te?^{n}de [te?^{n}de] 'to complete'
(not *[t^{n}ẽ?^{n}de])
b. ta?nu [t^{n}ã?nũ] 'bent'

(15) ya?mi [ỹã?mĩ] 'sweet potato'

4 Syllable structure

4.1 Onsets

Syllables without onsets occur in either syllable of the word although they are much more common in the second syllable of a word than syllables with onsets. Syllables without onsets are less commonly found in the first syllable.

(16) a. o.o 'to be'
b. ko.o 'to be (FUT)'
c. ko?.o 'bowl'
d. u?.u 'to be in pain'

In fact, NM seems to exhibit a real preference for the second syllable of a word to be simply a vowel, often identical to the vowel in the stressed syllable.

Any consonant, including the glides /y/ and /w/ may occur as a one-segment onset, though the lateral /l/ and the labial consonants /m, v, w/ are not common.

(17)

	Word-initial		Word-medial	
a.	vixĩ	'cold'	savi	'rain'
b.	čaka	'fish'	suči	'child'
c.	kimi	'star'	xika	'to walk'
d.	k^{w}ita	'to be tired'	šik^{w}i	'milk'

e.	lasu	'braid'	vilu	'cat'
f.	minu	'an herb'	kimi	'star'
g.	nune	'to have time'	xani	'to plow'
h.	ⁿdawa	'to fly'	xa?ⁿde	'to cut'
i.	savi	'rain'	kisɨ	'to sleep'
j.	šini	'head'	kaše	'to sneeze'
k.	taka	'nest'	ita	'flower'
l.	wa?a	'well'	ⁿdawa	'to fly'
m.	xinu	'to run'	saxɨ	'nephew'
n.	yata	'back'	kayu	'to get burned'

Two-segment onsets occur only in the stressed syllable of the word. There are two types of branching onsets. The first consists of a consonant followed by a glide. The following consonants are the only ones which have not been found before a glide: the lateral /l/ and the labial consonants /kʷ, v, m, w/.

(18)	/Cy/		/Cw/	
a.	tyu#šini	'small star'	twa?a	'smart'
b.	kyuu	'day'		
c.	ⁿdya?a	'to see'	ⁿdwa?a	'canyon'
d.	čyuũ	'chicken'		
e.	xyoo	'griddle'		
f.	sya?a	'to pass'	swa?a	'thus'
g.	šya?ã	'a bird'		
h.			nwaa	'why?'
i.			ywa?a	'snow'

We treat /kʷ/ as a single segment, rather than a consonant plus a glide, even though it would fit into the pattern shown above. One reason for this is that, although glides occur with many consonant phonemes, there are only a few examples of each. On the contrary, /kʷ/ is very common in all types of words.

Another reason for treating both /kʷ/ and /ⁿd/ as units in the syllable structure rather than as sequences is that they both can occur in the onset position of the second syllable of a word. Branching onsets do not occur in such syllables.

(19)	a.	šikʷɨ	'milk'
	b.	ⁿdakʷɨ	'strong'
	c.	koⁿdo	'knee'
	d.	šeⁿdu	'navel'

The second type of branching onset in uninflected words is a strident fricative (/š/ or /s/) followed by a stop:

š^{n}diki 'cattle'. It is not very common in monomorphemic words, but such onsets are common as a result of the affixation of the causative morpheme /s-/ and its allomorph /š-/. Branching onsets with /s/ or /š/ include the clusters /š^{n}d/ and /sk^{w}/. Our analysis of /^{n}d/ and /k^{w}/ as units permits these clusters.[7]

Given that most words have only two syllables and that branching onsets occur only in the first syllable, the following patterns are the only possibilities and they are all attested:

(20)	V.V	oo	'to be'
	V.CV	ita	'flower'
	CV.V	kee	'to come'
	CV.CV	kaku	'to be born'
	CCV.V	kyuu	'day'
	CCV.CV	syawa	'to irrigate'

Thus the simple word in NM may be described by the following formula:

(21) (C)(C)V.(C)V

The pattern CV.CV is by far the most common, and CCV.CV is quite rare.

4.2 Rhymes

All rhymes in Mixtec consist of a single vowel. Any of the vowels may occur in both stressed and unstressed syllables, although /o/ and especially /ó/ do not occur as fre-quently as other vowels.[8] As already stated, checked vowels occur underlyingly only in the stressed syllable,[9] and nasalized vowels only in the final syllable underlyingly.

5 Verb morphology

The majority of NM verbs only occur in two tenses, which we refer to here as the future and the nonfuture. In this section we account for the allomorphy of these prefixes. We assume the strongest hypothesis, namely, that each prefix has primarily one underlying form. There may be some genuine suppletive allomorphy, as we mention below. Allomorphy is shown to be dependent largely upon (a) the number of syllables in the verb stem, and (b) the type of initial segment of the stem--vowel, nasal consonant, or oral consonant.

The paradigm in Table 1 gives two examples from each pattern group, as well as a count of the number of verbs which illustrate that pattern from a random corpus of 200 verbs. Morpheme breaks are indicated. (In the case where an epenthetic vowel is involved, the break is indicated before this vowel.)

We posit the underlying form /k^W-/ for the future morpheme for all verbs, with the exception of a single verb which has the allomorph /č-/: **čiʔi** '(the wind) will begin to blow'. We do not discuss this allomorph further. The nonfuture morpheme is usually /xi-/, but it has the (presumably) suppletive allomorph /y-/ for a small arbitrary class of verbs. We also do not discuss these allomorphs further. A relatively small number of rules is necessary to account for the allomorphy that is exhibited by the forms in Table 1.

5.1 Nonfuture tense prefix

We begin by looking at the nonfuture morpheme /xi-/. The vowel is deleted before any polysyllabic root, whether they are consonant-initial or vowel-initial. (All vowel-initial roots are polysyllabic.)

(22) /i/-Deletion: $i \dashrightarrow \emptyset \;/\; \_\_\_ + C_0 V C_0 V$

Thus underlying /xi-asi/ becomes **xasi** 'to close' by /i/-Deletion. The derivation of **saka** 'to knead' (NF) from underlying /xi-saka/ involves /i/-Deletion and also the loss of the /x/. A rule such as Consonant Deletion could be stated to account for this loss.

(23) Consonant Deletion: $\underset{[-cor]}{C} \dashrightarrow \emptyset \;/\; \_\_\_ + C$

However, this rule would duplicate unnecessarily the constraints already stated with respect to syllable structure. We assume, following Harris 1983, that segments which cannot be incorporated into the syllable structure by the end of the derivation undergo the Erasure Convention, which is given below.

TABLE 1: Sample Verbs

(Grouping refers to first segment of root.)

Future		Nonfuture	# in Corpus of 200 Verbs
(Multisyllabic, oral consonant)			136
kani?nu	'to visit'	kani?nu	
^{n}dakate	'to wash'	^{n}dakate	
(Multisyllabic, nasal consonant)			17
ko-nee	'to carry'	nee	
ko-nini	'to hear'	nini	
(Monosyllabic, oral consonant)			11
k-aka	'to walk'	xi-ka	
k-oto	'to take care of'	xi-to	
(Monosyllabic, nasal consonant)			4
ku-ni	'to know'	xi-ni	
ku-nu	'to run'	xi-nu	
(Vowel, Class I)			19
k-asi	'to close'	x-asi	
k-uya?ã	'to give'	x-uya?ã	
(Vowel, Class II)			8
k^w-aã	'to buy'	x-aã	
k^w-iso	'to carry'	x-iso	
(Vowel, Class III)			5
k-uni	'to embrace'	y-uni	
č-i?i	'to begin (wind)'	y-i?i	

(24) Erasure Convention: Segments not incorporated into syllable structure at the end of a derivation are erased (Harris 1983:35).

(Certain assumptions have to be made so that this convention will delete the correct segment; see Harris 1983.) Thus the Consonant Deletion rule is superfluous. We note, however, that the Erasure Convention must apply in such a way that underlying /k^w-s#k^w-^{n}da?u/ (F-CAUS#F-to be poor) surfaces as **s^{n}da?u** and not as **^{n}da?u**. That is, although the underlying /s/ cannot be syllabified when the following /k^w/ is present, the erasure of the /k^w/ allows the /s/ to form an onset with /^{n}d/.

Thus underlying /xi-saka/ becomes intermediate /x-saka/ by /i/-Deletion and surface **saka** by the Erasure Convention. Because most roots are polysyllabic and consonant-initial, in most cases /xi-/ does not appear overtly in the word. Although the analysis of tone is incomplete, the facts seem to provide further evidence for this derivation. Briefly, /xi-/ carries a high tone which is not lost when the prefix is deleted, but is associated in such instances with the following syllable.

5.2 Future tense prefix

Since most Mixtec verb roots begin with an oral consonant, underlying /k^w-/ usually is deleted by the Erasure Convention. Thus underlying /k^w-saka/ becomes surface **saka** 'to knead' (future). This erasure is prevented from applying in the case of most nasal-initial roots by the following rule which vocalizes the labialization before nasals:

(25) Vocalization: $k^w \text{ --> } ku \;/\; \underline{\quad} \underset{[+nasal]}{C}$

Thus underlying /k^w-ni/ becomes **kuni** 'to embrace' by the Vocalization rule, and the Erasure Convention cannot apply. The phoneme /^{n}d/ usually functions as a non-nasal consonant with respect to this rule. Thus underlying /k^w-^{n}do?o/ becomes surface **^{n}do?o** 'to happen'. If this prefix /ku/ occurs in the antepenultimate syllable, as it does in the case of multisyllabic nasal roots, the /u/ lowers to /o/ by the following rule:

(26) Lowering: $u \text{ --> } o \;/\; \underline{\quad} + \$ \ \$$

Thus underlying /k^w-nini/ becomes intermediate /ku-nini/ by the Vocalization rule and surface trisyllabic

konini 'to hear' by the Lowering rule. The /u/ of **kuya?ã** 'to give' presumably does not lower since it is part of the multisyllabic (assumed to be compound) root /uya?ã/.

Before other consonant-initial roots, and most vowel-initial roots, the labialization of /k^{W}/ is lost. This is described by the following rule:

(27) Delabialization: $k^{W} \dashrightarrow k \ / \ \_\_\_ \ +$

Thus underlying /k^{W}-asi/ becomes **kasi** 'to close'. A group of vowel-initial roots (Class II in Table 1) must be marked as exceptions to Delabialization.

The Erasure Convention is prevented from applying in the case of monosyllabic oral consonant roots by the application of the following rule which inserts a vowel identical to the root vowel before the root:

(28) Vowel Epenthesis: $\emptyset \dashrightarrow V_{i} \ / \ C \ \_\_\_ \ [\ C \ V_{i}$

Thus underlying /k^{W}-ka/ becomes surface **kaka** 'to walk' with Delabialization applying either before or after Vowel Epenthesis.

There are verbs which must be marked as exceptional in one way or another. For example, there are two nasal consonant initial roots, /ni?i/ 'to find' and /no?o/ 'to go home', which must be marked [-Vocalization]. They pattern like oral consonant roots rather than as nasal consonant roots. Thus underlying /k^{W}-ni?i/ and /xi-ni?i/ have the identical surface form **ni?i**, rather than ***koni?i** and **ni?i**, and underlying /k^{W}-no?o/ and /xi-no?o/ both surface as **no?o**.

There are a few oral consonant roots, including a small minority of /^{n}d/-initial roots, that pattern like nasal consonant roots. These roots must be marked as positive exceptions to Vocalization in that the rule applies although the consonant is not nasal.

(29)	a.	koteku	'to be alive'
	b.	kokuu	'to be'
	c.	kosiki	'to play'
	d.	kotuu	'to lay down'
	e.	koku?u	'to become sick'
	f.	ko^{n}detu	'to wait'
	g.	ko^{n}dee	'to begin'
	h.	ko^{n}dikí	'to follow'
	i.	ko^{n}dita	'to owe'

j. ko^{n}dya?a 'to see'
k. ko^{n}deka 'to take'

A few even more irregular verbs remain. These are listed in Table 2 with tentative morpheme breaks, since they appear to exhibit two more (suppletive) allomorphs for the nonfuture morpheme: /ka-/ and /kɨ-/. The major irregularities are found in the nonfuture form.

TABLE 2: Irregular Verbs

Future		Nonfuture
k-o?o	'to drink'	x-i?i
k-oo	'to be'	oo
ku-sú	'to sleep'	ki-ší
ko-siki	'to play'	ka-siki
ko-tuu	'to lay down'	ka-tuu
k^w-iso	'to carry'	x-iso
ku-te?ẽ	'to get dirty'	ka-te?ẽ

The rule orderings for the rules discussed above are given below:

(30) Vocalization --
feeds Lowering, bleeds Vowel Epenthesis and Delabialization
Vowel Epenthesis
Stress Assignment
/i/-Deletion
Lowering
Delabialization

5.3 Other prefixes

Other prefixes which must be discussed in the morphology of the verb are the repetitive morpheme /^{n}da-/ and the causative morpheme /s-/. The causative morpheme /s-/ has an allomorph /ł-/ which occurs before non-velar consonantal consonants.

(31) Palatalization: s --> š / ___ + C [+cons] [-back]

$$s \rightarrow š \;/\; \_\_\_ + \begin{matrix} C \\ [+\text{cons}] \\ [-\text{back}] \end{matrix}$$

(32) a. skaka 'to drive'
b. sk^{w}iso 'to carry'
c. syu?u 'to frighten'

d.	š^{n}da?u	'to cheat'
e.	ščyo?o	'to cook'
f.	štee	'to plant'

Palatalization applies after the Erasure Convention deletes unincorporated segments since underlyingly neither /nda-/ nor /s-/ occur directly affixed to a root; the root must be preceded by the future tense prefix, as shown in the following formula:

(33) $$[\text{TENSE} + \left\{\begin{matrix}\text{REP}\\ \text{CAUS}\end{matrix}\right\} [\text{FUT} + \text{V}]]$$

However, due to the morphophonemic rules shown above, this is not always obvious. Derived verb stems with /^{n}da-/ pattern like oral-initial roots in the same way as do non-derived /^{n}d/-initial roots. The derivations of some verb forms with /^{n}da-/ and /s-/ are given below, showing the various rules operating. (The capital letters indicate stressed vowels.)

(34)

	NF-REP-F-to know	F-REP-F-to see
UF	/xi-^{n}da#k^{w}-ni/	/k^{w}-^{n}da#k^{w}-^{n}dya?a/*
Voc	xi-^{n}da-ku-ni	k^{w}-^{n}da-ku-^{n}dya?a
Epen	———————	———————
Str	xi-^{n}da-kU-ni	k^{w}-^{n}da-ku-^{n}dyA?a
i-Del	x-^{n}da-kU-ni	———————
Low	———————	k^{w}-^{n}da-ko-^{n}dyA?a
Delab	———————	k-^{n}da-ko-^{n}dyA?a
Eras	^{n}da-kU-ni	^{n}da-ko-^{n}dyA?a
Palat	———————	———————
SF	^{n}dakUni	^{n}dako^{n}dyA?a
	'to recognize'	'to review'

*Positive exception to Vocalization

	F-CAUS-F-to be poor	F-REP-F-to do
UF	/#k^{w}-a?^{n}di/	/k^{w}-^{n}da#k^{w}-sa?a/
Voc	———————	———————
Epen	———————	———————
Str	k^{w}-s-k^{w}-^{n}dA?u	k^{w}-^{n}da-k^{w}-sA?a
i-Del	———————	———————
Low	———————	———————
Delab	k-s-k-^{n}dA?u	k-^{n}da-k-sA?a
Eras	s-^{n}dA?u	^{n}da-sA?a
Palat	š-ndA?u	———————
SF	š^{n}dA?u	^{n}dasA?a
	'to cheat'	'to fix'
	NF-CAUS-F-to walk	F-CAUS-F-to kill

UF	/xi-s#k^{w}-ka/	/k^{w}-s#k^{w}-^{n}da?u/
Voc	____________	____________
Epen	xi-s-k^{w}-aka	____________
Str	xi-s-k^{w}-Aka	k^{w}-s-k^{w}-A?^{n}di
i-Del	x-s-k^{w}-Aka	____________
Low	____________	____________
Delab	x-s-k-Aka	k-s-k-A?^{n}di
Eras	s-k-Aka	s-k-A?^{n}di
Palat	____________	____________
SF	skAka	skA?^{n}di
	'to drive'	'to explode'

	NF-REP-F-CAUS-F-to carry	F-REP-F-CAUS-F-to put
UF	/xi-^{n}da#k^{w}-s#k^{w}-iso/	/k^{w}-^{n}da#k^{w}-s#k^{w}-tee/
Voc	________________	________________
Epen	________________	________________
Str	xi-^{n}da-k^{w}-s-k^{w}-Iso	k^{w}-^{n}da-k^{w}-s-k^{w}-tEe
i-Del	x-^{n}da-k^{w}-s-k^{w}-Iso	________________
Low	________________	________________
Delab	x-^{n}da-k-s-k^{w}-Iso	k-^{n}da-k-s-k-tEe
Eras	^{n}da-s-k^{w}-Iso	^{n}da-s-tEe
Palat	________________	^{n}da-š-tEe
SF	^{n}dask^{w}Iso	^{n}daštEe
	'to load (a gun)'	'to transplant'

6 Compound words

Some compound words in NM involve the juxtaposition of two lexical morphemes. If the first of these morphemes has a second syllable consisting of only a vowel, that vowel is usually dropped. We refer to the resulting form as the truncated stem. The truncated word loses its stress; penultimate word stress in the compound is maintained. Checking is lost in a truncated stem since only a stressed syllable may be checked.

(35) Truncation: V ---> Ø / V ___ # $ $

(36) ve?e 'house' + kaa 'metal' => ve#kaa 'jail'
kuu 'to be' + twa?a 'smart' => ku#twa?a 'to learn'
taa 'to give' + nuu 'awhile' => ta#nuu 'to lend'
kaa 'to eat' + šini 'evening' => ka#šini 'to eat supper'

As when the repetitive and causative prefixes are added, the non-truncated, second verb of a compound must be in the future form.

(37) kii 'to come' (F) + koto 'to visit' (F) =>
ki#koto 'to come to visit' (F)

kii 'to come' (F) + koⁿdya?a 'to see' (F) =>
ki#koⁿdya?a 'to come to see' (F)

xĩĩ 'with' + ki?ĩ 'to go' (F) =>
xĩ#ki?ĩ 'to take' (NFUT)

It should be noted that in the last example the loss of the final vowel in the truncated form does not include the loss of nasalization, which may be evidence supporting an autosegmental analysis of nasalization in Mixtec.

Adjectives and adverbs may be emphasized through reduplication of the first syllable of the word, forming a type of compound.

(38)	ⁿdi?i	'blue'	ⁿdi#ⁿdi?i	'very blue'
	kʷii	'green'	kʷi#kʷii	'very green'
	kʷe?e	'bad'	kʷe#kʷe?e	'very bad'

Another type of compound word is formed by combining an irregularly shortened form of a classifying noun with another noun.

(39)

na-	(nana 'mother')	+ šiso 'in-law'	'mother-in-law'
ndi-	(kiti 'animal')	+ xako	'opossum'
ta-	(tata 'father')	+ šiso 'in-law'	'father-in-law'
te-	(tee 'man')	+ tyuũ 'task'	'authority'
te-	(ⁿdute 'water')	+ ya?a 'hot pepper'	'hot pepper sauce'
ti-	(kiti 'animal')	+ nana	'tomato'
tũ-	(yutũ 'tree')	+ yuxa 'resin'	'pine tree'
yã-	(ya?ã 'woman')	+ si?i 'female'	'wife'

7 Motion verbs

A small class of motion verbs exists which occurs with the plural word **kʷei**, rather than the preposed plural clitic **ka**. This plural marker **kʷei** has the penultimate stress pattern of a normal word. The plural form of these motion verbs resembles a compound word, formed by applying the Truncation rule above.

(40)

	singular	plural
'to return'	kii	ki#k^{w}ei
'to go home'	no?o	no#k^{w}ei
'to arrive'	^{n}daxaa	^{n}daxa#k^{w}ei
'to get up'	^{n}dota	^{n}dota#k^{w}ei
'to enter'	kivi	kivi#k^{w}ei

A few verbs deviate from this pattern of plural-forming, and are irregular in other ways as well.

(41)

	'to lay down'	'to look like'	'to go'
FUTURE			
sing.	kotuu		ki?ĩ
pl.	kondwei		xĩkwei
NONFUT.			
sing.	katuu	kaa	kwa?a
pl.	kandwei	kanda	kwãkwei
PAST			
sing.	katuu		xa?ã
pl.	kandwei		xãkwei

8 Clitics

When vowel-initial bound pronouns (namely, /-i/ '3rd person, child' and /-o/ '1st person inclusive') occur after a verb or noun, the following rule applies:

(42) Elision: $\acute{V} \text{ ---> } \emptyset \;/\; V(?) \;\underline{\quad}\;]X_0 \;\; V$

Although the final vowel is elided, the nasalization feature of this vowel remains and is reassociated with the clitic vowel. Therefore either nasalization must spread before Elision applies or this feature must be treated autosegmentally.

(43)

word		word + /-i/	word + /-o/
ta?a	'relative'	ta?#i	ta?#o
staa	'tortilla'	sta#i	sta#o
sa?a	'to do, make'	sa?#i	sa?#o
nee	'to carry'	ne#i	ne#o
no?o	'to go home'	no?#i	no?#o
ki?ĩ	'to go'	ki?#ĩ	ki?#õ

The two clitics /-i/ and /-o/ may be added to the same word, yielding a string of three vowels.

(44) ka skuči#o#i
pl bathe-1incl-3,child
'we are bathing him'

Words ending in CV syllables do not undergo Elision but rather the following rule:

(45) Degemination: $V_i\ ]X_0\ V_i$ --> V_i

This is illustrated by the following forms which are underlined in the following paradigm:

(46)	word		word + /-i/	word + /-o/
	^{n}datatu	'to rest'	^{n}datatu#i	^{n}datatu#o
	yiki	'bone'	**yiki**	yiki#o
	kutoo	'to like'	kuto#i	**kutoo**

Some words ending V?V must be marked as exceptions to Elision (but not to Degemination).

(47)	word		word + /-i/	word + /-o/
	^{n}diši?u	'goat'	^{n}diši?u#i	^{n}diši?u#o
	či?i	'to plant'	či?#i	či?i#o
	xaše?ĩ	'to eat'	xaše?#ĩ	xaše?ĩ#õ
	ti^{n}do?o	'pitcher'	ti^{n}do?o#i	ti^{n}do?o
	ta?i	'squash plant'	ta?#i	ta?i#o

Deictic particles **yuũ** 'there' and **ya?a** 'here' have abbreviated clitic forms **-ũ** and **-a** which may be appended to nouns. The same rules of Elision and Degemination as for other clitics apply.

(48) tee 'man' + -ũ => te#ũ
kiti 'animal' + -ũ => kiti#ũ
kiti 'animal' + -a => kiti#a

Notes

* Ñumí Mixtec is spoken in the municipality of San Juan Ñumí and surrounding towns in the district of Tlaxiaco, Oaxaca, Mexico. The data are from San Antonio Nduaxico Ñumí, a town about one mile from San Juan. Gittlen, under the auspices of the Summer Institute of Linguistics, did field work there from March to September of 1982 and has continued her studies on the language since that time.

1. Glosses of verbs are in the infinitive although there is no infinitive form in Mixtec. Verbs will be marked FUT (future) or NFUT (nonfuture) when the segmental representation clearly indicates one or the other.

2. A few irregular nasal consonant-initial verb roots have more forms, as shown by the following verb **nini** 'to hear'.

Future tense

	incompletive	potential
sing.	ko konini	konini
plural	vi ko konini	vi konini

Nonfuture tense

	continuative	completive	habitual
sing.	nini	ni xonini	ni xo xonini
plural	ka nini	ni ka xonini	ni ka xo xonini

The incompletive aspect marker **ko** (apparently derived from **koo** 'to be' (FUT)) and the habitual aspect marker **xo** can occur with any verb. The prefix /ko-/ is an allomorph of the future tense prefix /k^w-/, as derived by the rules given below. The prefix /xo-/ may be a remnant of a more regular nonfuture form of the verb 'to be' than is now used in the language.

3. Specifically, these rare phonemes occur only in the following morphemes: /d/ in the bound third person pronouns **-de** 'male adult' and **-de** 'liquid'; /ř/ in the related independent and bound pronouns **řu?ú** and **-ří** '1st person, addressing a child', and **řo?ó** and **-řó** '2nd person child'.

In this discussion of Mixtec phonology we purposefully ignore consonants which occur only in Spanish loan words, such as the bilabial fricative.

4. We have found three exceptions to this allophonic rule: **yuṹ** [žṹṹ] 'there, that'; **yoʔó** [žóʔó] 'full form of first person inclusive pronoun' pers incl pronoun' **yaa** [yaa] ~ [žaa] and its clitic form **-ya** [-ya] ~ [-ža] 'spirit, saint, god', which exhibit free fluctuation.

5. The following are the only words known to have unpredictably nasalized vowels in the stressed syllable:

^{n}dṍta	'to stand up'
^{n}dṍta	'to resist'
^{n}dṍso	'to fall down'
kẽ́ta	'to reach, to arrive'
kã́ta	'to jump'
ti#čĩ́ki	'wrinkled'
ti#xĩ̀xa	'hornet'

6. Although Feinstein 1979 has argued that prenasalized consonants should be treated universally as consonant clusters, our data support the opposite viewpoint. In NM /^{n}d/ does not pattern as a nasal + obstruent within words (nor, as will be seen, in the morphology), which would seem to be problematic for this analysis.

7. Glottal stop is not a consonant in our analysis and therefore does not fill an onset slot in words like [wa.ʔa] 'well', [ʔoo] 'to be', and [xa.ʔ^{n}de] 'to cut'. In the last example, the analysis of the glottal as a syllable-initial consonant would require a new type of consonant cluster, and would have to be permitted in unstressed syllables as well.

8. Spanish loan words ending in /o/ are almost invariably altered to end in /u/.

k^{w}ẽ́tu	'story'	(Sp. **cuento** /kwento/)
sapatu	'shoe'	(Sp. **zapato** /sapato/)

9. Alternatively, it might be claimed (as previous Mixtec analyses have) that glottal stop is a phoneme which is extremely restricted. It would occur in NM only as the optional right member of a branching rhyme in a stressed syllable.

References

Alexander, Ruth Mary. 1980. Gramática mixteca de Atatlahuca. MexicoCity: InstitutoLingüistico de Verano.

Bradley, C. Henry. 1965. A linguistic sketch of Mixteco of Jicaltepec. Norman,OK: SummerInstituteof Linguistics, 25.

Feinstein, M. 1979. Prenasalization and Syllable Structure. Linguistic Inquiry 10:279-326.

Harris,James W. 1983. Syllable structure and stress in Spanish: A nonlinearanalysis. LinguisticInquiry Monograph 8. Cambridge: MIT Press.

Hunter, Georgia and Eunice V. Pike. 1969. The phonology and tone sandhi of Molinos Mixtec. Linguistics 47:24-40.

North, Joanne and Jana Shields. 1977. Silacayoapan Mixtec phonology. Studies in Otomanguean Phonology, W. Merrifield, ed., 21-33. Dallas: Summer Institute of Linguistics and University of Texas, Arlington.

Pankratz, Leo and Eunice V. Pike. 1967. Phonology and morphotonemics of Ayutla Mixtec. IJAL 33:287-99.

Pike, Eunice V. and John Cowan. 1967. Huajuapan Mixtec phonology and morphophonemics. Anthropological Linguistics 9:1-15.

Pike, Eunice V. and Thomas Ibach. 1978. The phonology of the Mixtepec dialect of Mixtec. Linguistic and literacy studies in honor of Archibald A. Hill, vol. II: Descriptive linguistics, 271-85. The Hague: Mouton.

Zylstra, Carol F. 1980. Phonology and morphophonemics of the Mixtec of Alacatlazala, Guerrero. Summer Institute of Linguistics, Mexico Workpapers. MexicoCity: Instituto Lingüistico de Verano.

FORTIS/LENIS CONSONANTS IN GUICHICOVI MIXE: A PRELIMINARY ACOUSTIC STUDY*

J. Albert Bickford

1 Introduction
  1.1 Articulatory background
  1.2 A possible analysis
2 Procedure
  2.1 Data collection
  2.2 Instrumental measurements
3 Results
  3.1 Amplitude
  3.2 Consonantal length
  3.3 Vowel length
4 Discussion
  4.1 The analysis of the fortis/lenis contrast
  4.2 Considerations for further study
  4.3 Comments and addresses

1. Introduction

This paper focuses on the fortis/lenis contrast in consonants in the Guichicovi dialect[1] of Mixe (Mixe-Zoquean; Oaxaca, Mexico), and uses instrumental data to support a proposed phonological analysis of this contrast. In the remainder of Sect. 1, I briefly survey prior work on this contrast, especially as it impacts on the claim that Mixe has three contrastive vowel lengths. I summarize the phonetic and phonological facts that were determined by ordinary techniques of phonetic transcription, and describe the phonological hypothesis for which I wanted confirmation from instrumental data. In Sect. 2, I describe the instrumental study, including the words tested, the recording procedures, and the measurements made. In Sect. 3, I present the results of this study. In Sect. 4, I discuss the implications of these results for Mixe and for phonological theory, and present some considerations for the design of a larger, more complete study.

1.1. Articulatory background

Hoogshagen (1959) first presented data from the Coatlán dialect of Mixe that appears to support an analysis of three phonemic vowel lengths. Especially convincing are minimal triplets such as the following (taken from Van Haitsma and Van Haitsma, 1976:9)

(1)	ˀoy	'although'
	ˀooy	'he went'
	ˀoooy	'very'

However, Norman Nordell (personal communication) later discovered that in the closely-related Guichicovi dialect, the difference between half-long and long is not contrastive. Rather, it is conditioned by a contrast in the final consonant, which he characterizes as fortis/lenis. More generally, half-long vowels always occur preceding fortis consonants, and full-long vowels preceding lenis consonants. Since the fortis/lenis contrast also occurs following short vowels, he regards it and not vowel length to be underlying; the variation in vowel length is thus conditioned. This contrast is subtle; it escaped his attention for many years. He speculates that something similar may be happening in Coatlán (based on informal auditory impressions of that dialect), although he has not been able to do any careful study. I have heard what appears to be the same contrast in the Mazatlán dialect, which is the third main Eastern dialect of Mixe.

Thus, an understanding of the fortis/lenis contrast in Guichicovi is important to the wider theoretical issue of whether a language can have three contrastive vowel lengths. It is the purpose of this paper to deepen our understanding of the fortis/lenis contrast and its effect on vowel length.

The fortis/lenis distinction is not easily stated in standard phonetic terms; prior to this study there seemed to be no single phonetic parameter which consistently correlated with the phonological contrast, and thus no easy way to characterize the contrast using standard phonological features. One important result of this study is the discovery of a consistent phonetic correlate to the fortis/lenis contrast which can therefore be posited as the "essence" of the phonological contrast.

As mentioned above, the phonetic length of a preceding underlyingly-long vowel is frequently an important phonetic cue for distinguishing fortis from lenis consonants. As the data below indicate, this is also true when the vowel is short, although this had not been noticed previously. This fact shows further the implausibility of the hypothesis that more than two vowel lengths are contrastive. If we were to claim that the fortis/lenis difference on consonants is conditioned by the length of the preceding vowel, we would need to posit _four_ contrastive vowel lengths (short, medium-short, medium-long, and long), which were further subclassified so that short and medium-long vowels formed a natural class which conditioned the following consonant to be fortis, while medium-short and long conditioned the following consonant to be lenis. It is much more reasonable to posit only two degrees of length, but to do so on both vowels and consonants.

The primary phonetic cue for the fortis/lenis contrast in obstruents is susceptibility to voicing.[2] The lenis versions of the consonantal obstruents p, t, k, c, and x are voiced in intervocalic position, whereas the fortis ones are always voiceless. At the end of a phrase, both fortis and lenis consonants are voiceless, but can generally be distinguished by the length of the preceding vowel. Utterance-initially, all obstruents are voiceless, but all are voiced when preceded by the personal proclitics **n-** '1st person' and **m-** '2nd person'. Nordell interprets this fact to mean that all word-initial obstruents are lenis. Thus [voice] could be regarded as the feature underlying the fortis/lenis contrast, assuming a rule which neutralizes the contrast in word-final and word-initial position.

However, [voice] does not work for sonorants, since both fortis and lenis sonorants are voiced. In all positions, fortis sonorants seem louder and longer than their lenis counterparts. As for h, the phonetic correlates of the contrast are also length and possibly amplitude. Previous to this study, both fortis and lenis versions of h were heard as voiceless, although this study shows that lenis h is subject to a certain amount of intervocalic voicing, like obstruents. For sonorants and h, then, consonantal length or gemination seemed like a better choice than voicing for representing the underlying distinction. One purpose of the present study was to measure the length and amplitude of these fortis/lenis pairs, and to determine which phonetic cue provides the most consistent basis for characterizing the fortis/lenis contrast using standard phonological features.

1.2. A possible analysis

Based on Nordell's description of the contrast, summarized above, I hypothesized that the underlying contrast in all cases was consonantal length, or more precisely, single (lenis) vs. geminate (fortis) consonants. This would directly account for the observed difference in length in sonorants and h. The greater amplitude of fortis consonants could be accounted for by a low level phonetic rule, or possibly by universal principles.

This analysis also provides a plausible account of the contrast in obstruents, assuming that single consonants are voiced between two voiced segments (i.e., either intervocalically, or following the nasal proclitics and preceding a vowel).

(2) Single consonant voicing
$C \rightarrow [+voice] \ / \ [+voice] \_\_\_ [+voice]$

Fortis consonants are immune to voicing because they are geminates.

Besides being very natural, this rule is supported by the fact that all obstruent clusters in Mixe are voiceless. Interestingly, all obstruent clusters in Nordell (n.d.) are written with an initial fortis consonant. This may have been Nordell's way of indicating that obstruent clusters, like fortis consonants, are immune to voicing.

To my knowledge, no consistent difference in length between fortis and lenis consonants had been observed prior to this study. Although the data below show that in careful speech, there is in fact a difference in length; it is possible that in fast speech the length contrast is reduced or eliminated, which might account for the failure to notice a significant phonetic contrast in length. In other words, it may be necessary to posit a rule such as the following:

(3) Degemination
$C_iC_i \rightarrow C_i$ (optional, fast speech)

This rule, if it is correct, would need to be ordered after the rules responsible for voicing single consonants and for adjusting the length of the preceding vowel. On the other hand, informal listening to recordings of casual speech seems to indicate that the difference in length is retained in casual speech, although I do not currently have the opportunity to check this out in detail.

Further evidence comes from Nordell's observation that all initial consonants are lenis. This follows automatically under my hypothesis from a rule that must be stated independently in the grammar of Guichicovi: words cannot begin with obstruent clusters.

Finally, it is very reasonable that long vowels would be shortened somewhat preceding double consonants. To test this fully, we should check to see if the same shortening occurs preceding a consonant cluster. Although Nordell (n.d.) does not mention this environment as one which causes shortening, it should be noted again that he writes obstruent clusters as if they contained fortis consonants.[3]

Thus, on purely phonological grounds, the hypothesis that fortis consonants are geminates has considerable attractiveness, both on the basis of language internal evidence, and on universal considerations of naturalness. The purposes of the current study were to seek instrumental confirmation of the phonetic description given above, to identify further phonetic detail which had not been noticed, and to test my hypothesis that the fortis/lenis contrast was essentially a matter of length.

Ultimately, Nordell and I would like to co-author a paper on this contrast, so that the current study also serves as a pilot study for our later work. Sect. 4 of this paper includes a discussion of the deficiencies of the work done so far, and how to correct them.

2. Procedure

2.1. Data collection

Nordell and I developed a set of six pairs of words which exhibited the fortis/lenis contrast. The words are given in (4) below. Four show contrast in intervocalic position. The other two exhibit the contrast in word-final position following a long vowel, and thus illustrate the two phonetic lengths of this vowel.

(4) Pairs of words used for studying fortis/lenis contrast

Consonant	Lenis	Fortis
p	kapɨk 'no (quot)'	kappɨk 'carry it (imp)'
h	kohɨp 'he should build'	kohhɨk 'build (imp, quot)'
n	tunɨp 'he should work'	tunnɨk 'work (imp, quot)'
y	huyɨp 'he should buy'	huyyɨk 'buy (imp, quot)'
t	peeet 'Peter'	peett 'he swept'
n	tuuun 'oblong, oval'	tuunn 'he worked'

In line with the hypothesis sketched above, I have written fortis consonants as geminates. For now, this can be regarded purely as a notational convenience. I have also distinguished the half-long vs. long vowels, even though this difference is not contrastive.

Most of these words are morphologically complex. This was considered unimportant, since the relevant phenomena are insensitive to the internal structure of words. Better pairs could probably be found, but these were deemed adequate for a pilot study.

Recordings of these words were made in April, 1984, in Mitla, Oaxaca, Mexico, with the help of Simón López, a native speaker of the Guichicovi dialect from the village of Mogoñé Viejo, who is in his late teens. Prior to recording each word, Nordell discussed it with him, using both Mixe and Spanish, to make sure that he understood which word he was expected to pronounce. Then we recorded Nordell saying the meaning of the target word in Spanish, and three tokens of López saying the target Mixe word, using careful pronunciation. Even with preparation, he at times said the wrong word, (in many cases because of confusion or ambiguity introduced by the use of Spanish), and sometimes Nordell was not satisfied that all three tokens were clear examples of the consonant under study. In these cases, we recorded an extra set of three tokens.

2.2. Instrumental measurements

Using this tape, I made spectrographs of each valid token. By "valid", I mean each instance of every word in our test list. I thus included those words which Nordell thought were not clearly pronounced, but discarded examples of other words which were recorded accidentally due to López's misunderstanding of our intent. In most cases, the unclear tokens yielded measurements comparable to the clear ones; apparently the particular phonetic cues which I was measuring were not the ones that Nordell relied on to determine clarity of pronunciation.

On each token, I measured the length of the fortis or lenis consonant, and the length of the preceding vowel. Then, for each word, I computed averages of the two measurements made on each token.

Determining the length of each segment was sometimes difficult because of the vague boundary that separates certain segments. Although nasals and voiceless obstruents displayed clear boundaries with adjacent vowels, the boundaries of other consonants were less clear. For example, regular glottal pulses on a vowel give way gradually to the random noise of /hh/, and /p/ (phonetically [b]) and y appear mostly as varied formant structures which show gradual transitions to the surrounding vowels.

I used a trace of average amplitude superimposed on the standard bar spectrogram to resolve this difficulty; somewhat arbitrarily, I considered a consonant to begin at the moment that the amplitude trace started decreasing at the end of the vowel. Although in some cases this "falling off" of the amplitude occurred significantly before voicing quit, it at least provided a consistent and precise criterion for measurement. Only in one case, y, did this not work, because there was no consistent drop in amplitude on this consonant. Instead, I relied on a relatively sudden shift in the frequency of the second formant at the beginning of the y. Data from this study about consonant and vowel length should not be lightly compared to data from other languages without taking into account the conventions used to measure these segments.

3. Results

3.1. Amplitude

Although I made amplitude traces for all the tokens, and had originally intended to measure amplitude of fortis and lenis consonants, I did not do so, for two reasons. The first was that something in the amplification circuits of the spectrograph was varying with the amount of time the machine had been turned on. This could conceivably make absolute measurements based on the amplitude traces unreliable, although the traces were still adequate for determining the boundaries of consonants.

The second reason was that it soon became apparent that amplitude was not going to provide a consistent phonetic cue for the fortis/lenis contrast. In intervocalic obstruents, the lenis stops, which were voiced, were louder than the fortis stops, which being voiceless had zero amplitude. On the other hand, the instrumental data seemed to support the characterization of fortis sonorants as louder, but this was not true in all tokens. Finally, when both fortis and lenis consonants were voiceless, as in the case of word-final obstruents, the amplitude was zero in both cases. Thus even when there was a discernible difference in amplitude, sometimes lenis consonants and sometimes fortis consonants were the loudest in a pair.

Without phonetic consistency, it would be difficult to claim that amplitude was the underlying "essence" of the fortis/lenis contrast. Although amplitude may be useful as a phonetic cue of the fortis/lenis distinction for subclasses of consonants, it does not appear to provide a good basis for a characterization of the phonological nature of this contrast.

3.2. Consonantal length

The measurements of consonant length for intervocalic consonants are presented in (5).

(5) Average length (in milliseconds) of intervocalic consonants

Consonant	Lenis	Fortis
p	60	179
h	136	224
n	68	81
y	86	133

Clearly, in each case, the fortis consonant was longer than the lenis one. This, of course, was expected for sonorants and h, based on our auditory impressions of these consonants. In the case of n, it is possible that the difference in length is not statistically significant, but the others show a large difference which is almost certainly significant.

One of the surprises of this study was the striking contrast in length with the obstruent p, which had not been previously noticed.[4] After being alerted to the possibility that obstruents might also differ in length, this difference was clearly discernable by ear.

Another fact not noticed by ear was a difference in voicing on h comparable to that noted for obstruents. In lenis /h/ there were clear glottal pulses throughout, but in fortis /hh/, glottal pulses very quickly dropped off to zero. Thus intervocallic lenis h is voiced, like other obstruents. Up to this point, I have avoided referring to h as an obstruent, since its precise analysis in distinctive features (i.e. [+son] or [-son]) has not been crucial. This susceptibility to voicing is evidence that h is [-son].[5]

Another surprise was that it was possible to measure the length of t in final position, even though both fortis and lenis versions are voiceless. Most tokens exhibited a release at the end of the consonant; whenever this was apparent, I measured it. (Measurement of n was easy, since it is voiced.) The results are shown in (6).

(6) Average length (in milliseconds) of final consonants

Consonant	Lenis	Fortis
t	131	229
n	85	104

Again, fortis consonants are longer than lenis ones, although the difference for n may not be significant. This difference had not been previously noticed on obstruents, since the phonetic cue (a slight pop when the stop was released) has such a low amplitude.

3.3. Vowel Length

The measurements of vowel length for underlyingly long vowels are given in (7).

(7) Vowel length (in milliseconds) for long vowels preceding final fortis and lenis consonants

Following consonant	/___ Lenis	/___ Fortis
t	262	142
n	312	205

As expected, underlyingly long vowels are significantly shorter preceding a fortis consonant.

The third major surprise in the study was that this phonetic difference in length also occurs on phonemically short vowels, as illustrated in (8).

(8) Vowel length in milliseconds for short vowels preceding medial fortis and lenis consonants

Following consonant	/___ Lenis	/___ Fortis
p	138	108
h	110	91
n	118	92
y	100	82

This difference is somewhat subtler than that which occurs with long vowels; this partly explains why it was not noticed before.

4. Discussion

4.1. The analysis of the fortis/lenis contrast

In all six pairs of words, exemplifying all major classes of consonants in both positions where the contrast occurs, the fortis/lenis contrast correlates directly with a phonetic difference in length. This correlation appears even in situations where it was not previously noticed. Only for n is this difference subtle enough that it may not be statistically significant. Thus the hypothesis that fortis consonants are geminates is not only reasonable phonologically, it has a strong phonetic basis as well.

As noted throughout, the putative third contrastive length in Mixe vowels is in fact a conditioned phonetic variation, at least in the Guichicovi dialect. This variation is also apparent in a subtler form on underlying short vowels. Although the subtlety of this difference for short vowels might raise questions of its significance, the consistency of the results suggests that either 1) native speakers are in fact aware of this conditioned variation, 2) they automatically generalize the rule for long vowels so that it also applies on short vowels, or 3) the process of vowel shortening before a double consonant is a natural process in the sense of Stampe (1973), and therefore does not need to be learned.

At any rate, the length of the preceding vowel appears to be a consistent (and most likely statistically significant) phonetic cue for the underlying contrast in consonant length.

Although amplitude and voicing may be significant phonetic manifestations of the fortis/lenis contrast, they do not provide consistent correlations with the underlying contrast, since they vary depending on the type of consonant and its environment. Therefore, these variations are best analyzed as the result of conditioned variation.

The fact that the Mixe fortis/lenis contrast can be identified with underlying gemination suggests that other cases where the fortis/lenis distinction has been claimed could also be the same phenomenon. If so, the failure to identify the contrast immediately as one of gemination could be due to the operation of phonological rules such as voicing and spirantization of single consonants, and possibly subsequent degemination. For example, the description of the fortis/lenis contrast in Cajonos Zápotec given by Nellis and Hollenbach (1980) is very similar to what has been noted in Mixe. Although they argue (p. 103) against the possibility that fortis/lenis in Cajonos is a matter of gemination, they do not reject the possibility (p. 98) that it might be due to an underlying feature [long] on a single consonant. Similarly, McKinney (1984), using instrumental data, demonstrates that the fortis/lenis contrast in initial consonants in Jju (also known as Kaje, a Benue-Congo Plateau language of Nigeria) is consistently correlated with length; see his paper for details on the measurements involved.

4.2. Considerations for further study

Certain weaknesses in the pilot study are summarized here, so that the later complete study can avoid them; other design requirements for that study are also noted.

For a more complete study, a larger set of words should be used. Attempt should be made to find examples of the fortis/lenis contrast for all consonants and all positions, including palatalized consonants as well as nonpalatalized. Vowel length should also be studied before clear cases of consonant clusters, and the other six types of vowel nuclei should be exemplified.[6]

Some set of consistent, reliable, and most importantly, standard criteria should be developed for fixing the boundary between consonants and vowels for the purposes of measurement. The measurements taken by McKinney (1984) should be strongly considered.

In the current study, all three tokens of a word were recorded in sequence. A better approach, which controls for the effects of tiredness, list intonation, etc., would be to record each token separately in a frame, and to randomize the order of recording. This may require a literate subject. Another way to avoid this problem would be to record tokens in an alternating sequence, thus:

(9) huyɨp
huyyɨp
huyyɨp
huyɨp

The recording quality, although adequate for a pilot study, could have been improved. We recorded the words in a kitchen, so that there was a fair amount of noise in the background. It was difficult to control the recording level, because the microphone was hand-held and López had difficulty controlling the distance from it to his mouth. These two factors combined made it difficult to get a good spectrogram for some tokens. Since there are at least two professional quality recording studios in the Mitla area which are available for use in linguistic studies, these facilities should be employed in the next study, together with the assistance of professional recording technicians.

I had no clear notion of how much difference in length could be considered statistically significant. Some thought must therefore be given to statistical reliability.

At some point, it may be desirable to explore the relative importance of the various phonetic cues (consonant length, vowel length, voicing, and amplitude, and possibly others) for the perception of these words. This would involve synthesized speech presented to native speakers in some controlled fashion. However, such a study clearly must follow a more careful and complete determination of the relevant phonetic parameters.

Finally, in order to settle the issue of the number of contrastive vowel lengths in Mixe, it must be determined whether the fortis/lenis contrast is also present in the dialects of Coatlán and Mazatlán. Some spectrographic studies of these dialects may also be desirable.

4.3. Comments and addresses

Comments on this paper are very welcome. Until at least May 1986, I can be reached at UCSD, and can always be reached through SIL in Tucson.

Dept. of Linguistics, C-008
University of California, San Diego
La Jolla CA 92093

Summer Institute of Linguistics
Box 8987 CRB
Tucson AZ 85738-0987

Notes

*I wish to express my appreciation to the phonetics lab at the University of California, San Diego, for the use of the sound spectrograph; to Simón López, for his assistance in recording the list of test words; and to Margaret Langdon and Hu Mathews for comments on earlier drafts. Most importantly, this study would have been completely impossible without Norman Nordell, whose knowledge of Mixe is far more extensive than mine. Unfortunately, it has not been possible to obtain comments from him on the results or on the discussion below; the results of this paper must therefore be regarded as tentative. Due to the preliminary nature of this study, please do not make direct quotations without first consulting me. The inevitable errors are of course my own.

[1] Guichicovi Mixe is spoken in and near the municipio of San Juan Guichicovi, in eastern Oaxaca, Mexico, just north of Matías Romero. It has also been called Isthmus Mixe, Eastern Mixe, and Mixe of Mogoñé.

[2] Both fortis and lenis versions also are subject to palatalization, giving a four-way contrast at each point of articulation. However, palatalized consonants are not included in this study. For simplicity, I will often refer to a point of articulation with a single letter which is not surrounded by slashes; thus p in the discussion will stand for both the fortis and lenis bilabial stops. In transcriptions of lexical items, or within slashes, /p/ will represent only the lenis stop. Two points of articulation mentioned here will not be discussed further: c representing coronal affricates /ts, tts/ and x representing coronal fricatives /ṣ̌, ṣ̌ṣ̌/.

[3] To make the test completely valid, the cluster must be tested after unstressed syllables, since consonants are frequently strengthened after stressed syllables. (See the various strengthening rules in Nordell n.d.:44-61)

[4] At least, I had not noticed it, and I don't remember Nordell ever mentioning this to me.

[5] In one environment, the rule of intervocalic voicing is bled by another rule. An initial consonant cluster consisting of a nasal stop plus a (lenis) h coalesces to a voiceless nasal stop, rather than causing the h to become voiced. For example, m + huyɨp surfaces as [Muyɨp], where [M] is a voiceless bilabial nasal stop.

[6] These involve various combinations of short and long vowels with /h/, /ʔ/, and laryngealization.

References

Hoogshagen, Searle. 1959. Three contrastive vowel lengths in Mixe. Zeitschrift für Phonetik 12:111-15.

McKinney, Norris P. 1984. The Fortis feature in Jju (Kaje): An initial study. Studies in African Linguistics 15:177-188.

Nellis, Donald G. and Barbara E. Hollenbach. 1980. Fortis versus lenis in Cajonos Zápotec phonology. IJAL 46:92-105.

Nordell, Norman. n.d. El Mixe de San Juan Guichicovi. ms., Instituto Lingüístico de Verano, Mexico.

Stampe, David. 1973. A dissertation on natural phonology. Ph.D. dissertation, University of Chicago. In press: Garland, New York.

Van Haitsma, Julia Dieterman, and Willard Van Haitsma. 1976. A hierarchical sketch of Mixe as spoken in San José El Paraíso. SIL publications in linguistics and related fields 44. SIL (University of Oklahoma session), Norman, Oklahoma.

THE INFLECTIONAL/DERIVATIONAL DISTINCTION[1]

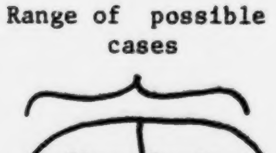
David Tuggy

1 Introduction
2 Prototypical inflection and derivation
  2.1 Productivity
  2.2 Predictability
  2.3 Meaning change
  2.4 Summary
3 Intermediate cases
4 Conclusion

1 Introduction

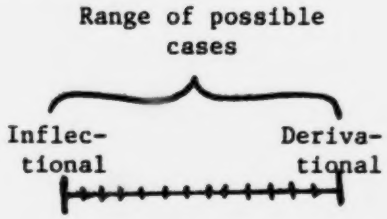
The traditional distinction between inflectional and derivational morphology is a useful one, but it is quite complex and difficult to apply to actual forms. With this, as with other linguistic distinctions, there is a tendency to assume that we have an absolute binary distinction, a dichotomy, a hard-and-fast line, with everything on one side purely inflectional and everything on the other side purely derivational. I do not believe the facts fit such a model.

I would claim that the distinction is better viewed as gradual or scalar along several parameters. In other words, rather than the dichotomy presented in 1.a, the scale of 1.b is a more accurate representation, with its possibility of a morpheme falling in between the two poles on the scale, of one morpheme being more inflectional or more derivational than another, without it being possible to unequivocally call some morphemes either inflectional or derivational. This notion of a scale or cline has affinities to Pike´s wave model, which in 1.c. is applied to this distinction.



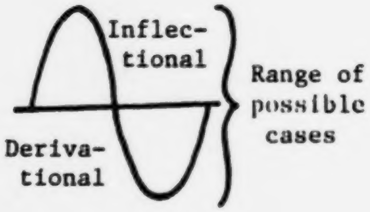
Range of possible cases
Inflec-tional
Deriva-tional

1.a



Range of possible cases
Inflec-tional
Deriva-tional

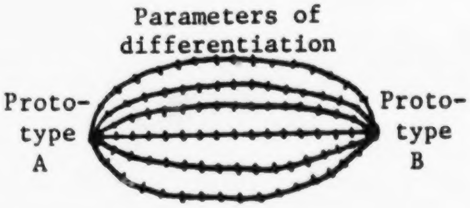
1.b



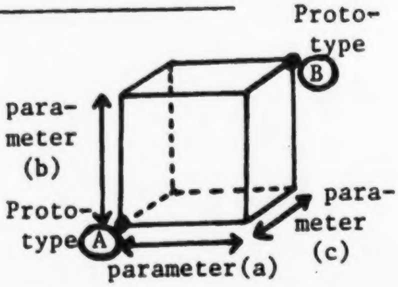
Inflec-tional
Deriva-tional
Range of possible cases

1.c

Such scalar distinctions are expected in what has been called the "prototype" model of categorization, which many cognitive psychologists believe more accurately represents normal human categorization than does the simple dichotomous model. Typically there will be several parameters along which contrasting categories will differ;[2] as a limiting case there may be only one: typically the distinctions along these parameters are scalar; again as a limiting case they may be binary or dichotomous. In 1.d is diagrammed a schema for such categorization: 1.e shows a cube-shaped model that can be employed when only three parameters are involved. We will use such diagrams in the following discussion, for I will be concentrating on three dimensions or parameters of the inflectional/derivational distinction.



1.d Prototype Model



1.e Three-Dimensional Prototype Model

2 Prototypical inflection and derivation

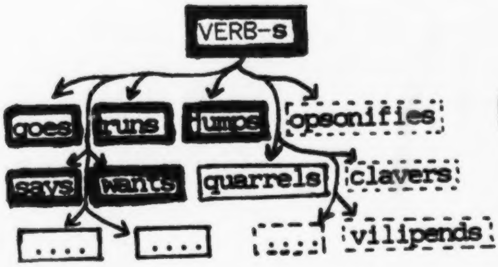
As a prototypical inflectional morpheme, let us take the third person singular present-tense marker **-s** in English, and as an example of a derivational morpheme the adjectivizing suffix **-some.** Both of them can occur, for instance, with the stem **quarrel.** I think everyone will agree that **quarrels** is an inflectional variant of **quarrel**—no one would seriously consider listing it in a dictionary, for instance. But everyone would list **quarrelsome.** What are the differences between them? I will sum them up under three heads: productivity, predictability, and meaning change.

2.1 Productivity

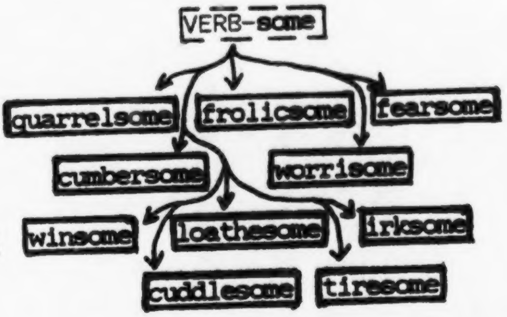
One big difference between -s and **-some** is this: **-s** can be, and in practice is, attached to virtually any English verb. Consider the following synonyms for **quarrel: fight, argue, disagree, brangle, wrangle, bicker, squabble. -s** can be attached to any of them, and the result is clearly English: **fights, argues, disagrees, brangles, wrangles, bickers, squabbles.** In contrast, **-some** sounds various degrees of bad, from odd to execrable, with all but **quarrel:**

fightsome, **arguesome**, **disagreesome**, **branglesome**, **wranglesome**, **bickersome**, **squabblesome**, and (Aaahh!) **quarrelsome**.

Depending on your linguistic theory this difference can be represented in varying ways. I like to express it as follows:[3] **-s** has associated with it a schematic VERB-s structure, firmly entrenched, through constant usage, in English speakers´ linguistic inventory, and constantly used to sanction novel, as well as established, structures. Individual instances of this structure, such as **quarrels**, may well also be established, but even if they weren´t they could and would readily be computed anew. This is diagrammed in 2.a: The established structure VERB-s sanctions both other established structures and novel (non-established) structures of the same type. (Degree of establishment or entrenchment is represented by continuity and thickness of the box enclosing the representation of a form; the sanctioning or schematic relationship is indicated by the arrow.)



2.a VERB-s Structures



2.b VERB-some Structures

In contrast, while **-some** may be claimed to have a VERB-**some** structure associated with it, that structure is probably not firmly entrenched, and it clearly is not commonly used to sanction novel structures. **Quarrelsome**, **fearsome**, **frolicsome**, **loathesome**, and a few other (I´ve thought of six) specific structures, whose commonality it represents, are reasonably claimed to be a good deal more firmly entrenched than it is. This situation is diagrammed in 2.b.

Productivity tends to correlate inversely with awareness that any particular productive usage is in fact novel. Once a new verb that has never been used in third person singular (perhaps **opsonify** or **vilipend**) is established, the novelty of a formation **opsonifies** or **vilipends** will pass unnoticed. In contrast, a formation

opsonifysome or **vilipendsome** (or **arguesome**, etc.) would command immediate attention.

In sum, (proto)typically inflectional morphological structures tend to be productive, smoothly sanctioning many structures, both established and novel; while (proto)typically derivational structures tend to be unproductive, sanctioning few established structures and fewer novel structures, and provoking a strongly conscious reaction when they do sanction novel forms.

2.2 Predictability

A related, but different, difference is predictability of function or semantic effect. **-s** has the same semantic effect every time it is used: it is used with a verb, that is a structure designating a process of some kind, and it specifies that the subject of that process is neither the speaker nor the hearer, but some other single entity adequately identified to both. (We will ignore here its present tense specification.) It is rigidly limited in how it functions, which permits it almost unlimited application: hearers of any novel form using it will automatically understand what is meant. (There is a parable here.)

-some, on the other hand, functions rather differently in different words. In **quarrelsome** it attaches to a verb, and describes some entity involved in the process designated by that verb as possessing qualities tending to the occurrence of the process. That entity may be a person or some aspect of a person (e.g. his attitude or tone of voice), but that person will, if the quarreling occurs, be involved in it: **quarrelsome** does not mean "provoking quarreling in others". The dishes seem to have qualities tending towards my children quarrelling when they have to do them, but I would not call the dishes **quarrelsome.** In **fearsome, -some** again attaches to a verb, and again it describes an entity as possessing qualities tending towards the occurrence of the process, but this time towards its occurrence in others. The **fearsome** thing does not fear, rather it causes others to fear.[4] If **fearsome** were really parallel to **quarrelsome**, it would mean the same as **fearful**, but it doesn't.[5] Other cases are even more different. **Wearisome** is much like **fearsome** except that **-some** here attaches to an adjective.[6] In **fulsome, -some** again attaches to an adjective, but now describes a process (such as praise) as being characterized by that adjective to an excessive degree. **Toothsome** has **-some** attaching to a noun, producing an adjective describing some other thing as having qualities tending to make one want to indulge with respect to it in a process crucially involving the thing designated by the noun stem. In **handsome** it attaches to—what? another noun?—and characterizes a thing as (quite unpredictably) pleasant and impressive to look at. There are almost as many functions for **-some** as there are words in which it occurs. This naturally affects productivity: any time a new structure using **-some** is encountered, the hearer must guess (he cannot know for certain) what is intended. Its flexibility severely limits its usefulness.

In sum, inflectional morphology tends to have predictable functions and semantic effects; derivational morphology tends to be act unpredictably.

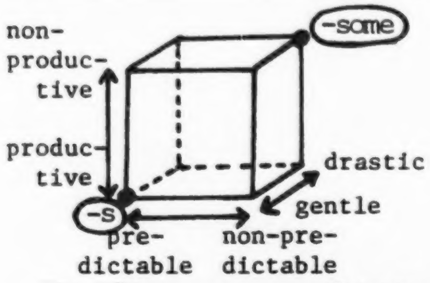
2.3 Meaning change

-s and **-some** differ also in the extent to which they change the semantics of the stem to which they are attached. Virtually all morphology changes meaning to some extent—otherwise we would not bother to use it. But there is a difference between minor adjustments to the meaning and major overhauls of it.

-s provides a minor adjustment, a slight nudge, to the meaning of **quarrel**, or of any other verb it attaches to. It simply provides a bit of information already half expected but not yet specified. **Quarrel** specifies a process in which one person engages in a certain type of combative activity with one or more others: **quarrels** differs from that only in that the identity of the protagonist is limited so as to exclude speaker, hearer, and plural entities. **-some**, on the other hand, performs a major overhaul. Instead of the process designated by **quarrel**, **quarrelsome** designates a quality, which may but also may not result in the occurrence of that process. **-s** fits neatly into the scheme provided by **quarrel**; **-some**, in contrast, grabs **quarrel** and stuffs it into its own scheme. The inflectional morpheme is gentle, making small semantic changes; the derivational morpheme is rough, making drastic changes.[7]

2.4 Summary

I am claiming, then, that, prototypically, inflectional morphology is productive where derivational morphology is non-productive; inflectional morphology is predictable where derivational morphology is not predictable, and that inflectional morphology makes only minor adjustments to the semantic structures it affects, whereas derivational morphology makes drastic changes to them.[8] We can diagram these three differences on a cube like that in 1.e. The vertical dimension we will use to represent the parameter of productivity, the left-right horizontal dimension to represent predictability, and the front-back dimension to represent the gentle vs. drastic semantic change induced in the stem by the morphology.



2.c Parameters of Differentiation between Inflectional and Derivational Morphology

3 Intermediate cases

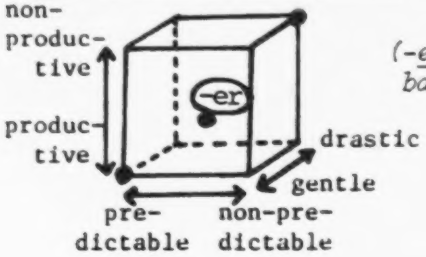
If, as I want to claim, these parameters are scalar rather than dichotomous, and if, in fact, all three are relevant, it ought to be possible to find intermediate cases, morphemes that would have to be located elsewhere than at the lower left corner of the cube like **-s** or the upper back corner like **-some**. They are not difficult to find.

Consider **-er**, for example. It, like **-s**, is very productive, and can be applied to almost every verb in the language. **Quarreler, arguer,** and **disagree-er** are not established words in my dialect of English, but I find them perfectly understandable; **fighter**, and to a lesser extent **squabbler** feel like they are already entrenched. Also, **-er**´s function is largely predictable: at least when it is being used productively it virtually always changes the designation of a form from that of a process to that of a thing, the subject of that process. However, that is not a gentle change, but a rather brusque one—nouns in **-er** are in fact what we call derived nouns, just as adjectives in **-some** are derived adjectives, whereas verbs in **-s** are not derived verbs, but inflected.

However, the degree of **-er**´s productivity is less than that of **-s**. **-er** is productive in that it may be applied to virtually any verb in English: however in practice it is not applied to so many as **-s** is. **-s** is applied so constantly to form so many novel forms that we are hardly aware when we are hearing or even producing a new one. With **-er**, in contrast, we do not so constantly hear and produce novel forms, and thus we are more aware of a new one when we do hear it.

Furthermore, note that **-er** is not entirely predictable, in several ways: (1) There are many firmly established constructions in which the nature of the thing designated by the VERB-**er** structure is specified further. E.g. a **computer** is no longer just any thing/person that computes; a **rocker** is not just any thing/person that rocks, nor a **fender** just any thing that fends. These forms have become semantically specialized in reasonable, but unpredictable ways. (No one fifty years ago could have predicted that a calculator could not be called a **computer**; if **-s** were to behave similarly we would have forms such as as **walks** meaning "(a woman) walks", or **runs** meaning "(a horse) runs".) In many cases this extra semantic material becomes so prominent that it takes over, displacing the process designated by the verb stem to the point where it figures only very marginally or figuratively if at all in the characterization of the resultant structure. Calling someone a **stinker** does not usually imply that he literally stinks, nor does a **hooker** literally hook. The **clinker** may clink, but that is surely not a very central facet of its meaning. (2) Sometimes **-er** combines with a verb but designates a thing other than the subject of the process designated by the verb. In **drawer** and **(all day) sucker** and perhaps **trailer** and **locker** it is the object; in **drainer** and **diner** (referring to a railroad car) and perhaps again **locker** it is the location; in **sweater, loafers,** and **pedal-pushers** it is clothing worn which tends to permit one to engage in the activity; in **dinner** and **supper** it is the occasion; in **ouster**, **prayer**, and **reminder** it is the action of the verb itself, nominalized; in **bender** (as in "go on a bender") it is not clear what it is. (3) **-er** often combines with stems that are not simple verbs or are not verbal at all: e.g. in **back-scratcher, pancake-turner, pencil-sharpener,** and **can-opener** it attaches to a verb with an incorporated object, even though the verb may not independently occur with such an object (***back-scratch, *pancake-turn**); in **do-good-er** it attaches to a more normal-looking verb + direct-object phrase; in **hum-dinger** it applies to an apparently complex stem that may or may not be verbal; in **grounder, beaker, saucer, 3-wheeler, header, facer, porker,** (ocean) **liner, SIL-er, heather, teenager, westerner, bummer, downer, insider, up-and-out-er** it applies to non-verbal stems. In all these cases its function or semantic effect is different and is not predictable (though it is reasonable to some extent). (Note that **-er**'s use with such stems is not unproductive, even though it is not as freely productive as the VERB-**er** Subject usage described above.)

Thus we can diagram the position of **-er** on our cube as in 3.a: it is productive like, though not as much as, **-s**, but it makes a drastic change, like **-some**, and though it is more predictable than **-some**, it is less predictable than **-s.**[9]



(*-er* *is on the back wall of the cube)*

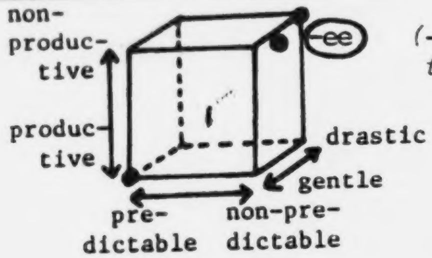
3.a -er

The suffix **-ee**, which we often think of as opposite to **-er** (on the basis of such pairs as **employer/employee** or **payer/payee**), provides another example of an intermediate case. I would judge that **-ee** is a good bit closer to the derivational terminus than is **-er.** In the first place, **-ee** is not nearly as productive. There are few verbs to which one cannot affix **-er**, but many to which it is difficult if not impossible to affix **-ee.** **-ee** prototypically designates the human direct object of a verb, which naturally reduces its range to transitive verbs which can take a human direct object. (That is why e.g. **plowee** is malformed: we do not plow humans.) However, even for verbs that take human objects, **-ee** is not freely productive. **hearee** sounds next to impossible, as do **appreciate-ee, fightee, love-ee, killee, remindee, rememberee, tease-ee** and any number of other forms. Yet **-ee** is not totally unproductive: new forms do turn up. While writing this paper, I found, in Time and Reader's Digest, the forms **tippee** and **fallee**, both new to me. Most of us have even made **-ee** forms up ourselves on occasion; we tend to be very aware when we are doing so, however. The list of forms with **-ee** is clearly not as small and closed as the list of forms with **-some.**

With regard to predictability, **-ee** is perhaps less predictable than **-er**, and not much more predictable than **-some.** Its semantic effect is normally to change a process (verb) to designate a (non-subject) human involved in that process, typically the direct or indirect object. Even in this usage it often has extra semantic baggage with it; for instance a **referee** is certainly not just anyone referred to, but rather the person referred to in a very specialized sort of situation, and similarly a **committee** is not just anyone to whom something is committed. Moreover, it has other usages. In **amputee** it designates not the direct or indirect object, but the person whose body-part was the direct object amputated.[10] More often it is used (especially in business-ese or militar-ese) to designate the subject of an intransitive verb (often a verb which might be construed as a verb of suffering). **Attendee, standee, escapee**, and Relational Grammar's **ascendee** are examples that come to mind. (That last example designates a non-human entity, which is another unpredictable twist.) And in one case, **consultee**, the subject of a transitive verb rather than its object is designated, the person who consults someone else rather than the person consulted (who is called, just as illogically, the **consultant**). **Addressee** can be quite plausibly analyzed as involving the noun **address** rather than the verb; at least for me the prototypical meaning of **addressee** is "person whose name is in the address part of the envelope" (or perhaps "person to whom the letter is addressed") rather than "person addressed". Similarly **conferee** is probably to be derived from the noun **conference** rather than the verb **confer**.

The brusqueness of the semantic shift produced by **-ee** is comparable to that produced by **-er**; shifting from a process (typically) to designate a person involved fairly centrally in that process.

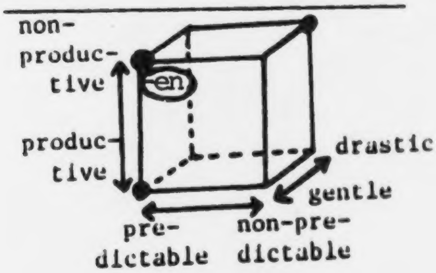
Thus we could diagram the place of **-ee** on our cube as in 3.b below.



(-ee is also on the back wall of the cube)

3.b -ee

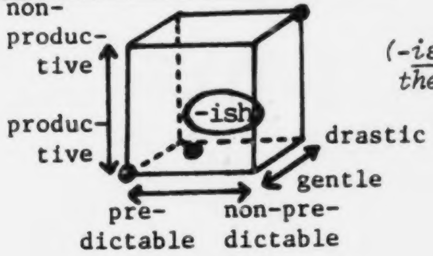
As a further example, consider the plural suffix **-en** as in **oxen, children, brethren.** It is quite predictable in its functioning, being suffixed to a noun stem and changing the designation from a single thing (the meaning of the noun stem) to a group of such things. This change is not very drastic, and in fact if linguists were asked for a class of typical inflectional morphemes, many would suggest plurals, which can be defined as those morphemes which accomplish exactly this semantic change. In these two aspects, then, **-en** is clearly like an inflectional morpheme. However, it is not at all productive, applying only to one relatively infrequent stem (**ox**) and in one archaistic form (**brethren**), in one frequent form which is irregular in other respects as well (**children**), and perhaps (though this is a bit farfetched), in two other very common irregular forms (**men** and **women**) and their derivatives. Thus by the third parameter **-en** is clearly like derivational morphemes. **-en**'s position can be diagrammed as in 3.c.



3.c -en

As a final example, consider the suffix **-ish.** It is quite productive: it typically suffixes to adjectives, and can be readily suffixed to the vast majority of the adjectives in English. (Adjectives that designate qualities not readily construable as matters of degree resist the suffixation to some degree: **pregnantish** sounds odd to me, but not impossible.) As usual, some **-ish** constructions are well-established in their own right (**reddish, smallish**), others are more marginal (**blackish, tightish, large-ish**), others feel virtually novel (**chartreuse-ish**). Not only is **-ish** productive, like a prototypical inflectional morpheme, but it accomplishes only a small shift in the meaning: it takes an adjective and adjusts it so it designates the same quality as before, but in attenuated form; **reddish** means "red, but not strongly red."

However, **-ish** also turns up on a large number of noun stems, converting them to adjectives denoting the possession of a quality associated with the noun.[11] This usage is both less productive (though still productive) and semantically more brusque. Furthermore, it often has extra semantic baggage associated with it. For instance, **-ish** often produces negative terms, even when the noun stem it attaches to carries no such implication: **childish, mannish,** and **womanish** are examples; in **foolish, brutish,** and **slavish** the negative overtones may come from the stem instead or as well. In my speech at least, although **childish** is a negative term, **girlish** can be either negative or neutral, and **boyish** has positive connotations. This is a symptom of a larger problem, namely that there is variation (and therefore unpredictability) with respect to which quality of the noun is the basis for the adjectivalization. It may be a virtually defining characteristic (**foolish**) or a rather peripheral one (**old-maid-ish**). Then too there are cases where **-ish** attaches to stems that are difficult or impossible to identify: what is **brack-** in **brackish**, or **squeam-** in **squeamish**? Thus **-ish** is less predictable in its functioning than, say, **-s.**



(-ish is slightly above the bottom and to the front of the cube)

3.d -ish

4 Conclusion

Dozens of other examples could have been cited from English, or from any other language that I know anything about. If they are plotted on a cube like those in 1.e, 2.c, or 3.a-d, there will be some tendency to cluster towards the two poles of inflectionality and derivationality, but there will be examples at many intermediate points, with no clear boundary separating the one cluster from the other.

To me, the conclusion is clear: the distinction between inflectional and derivational morphology is not a sharp dichotomy. Prototypical examples from the two categories differ along at least three important parameters, each of which is scalar rather than dichotomous in nature. To the degree that a morpheme or morphological construction is productive, predictable, and mild in the semantic effects it produces, to that extent it is inflectional; to the extent that it is unproductive, unpredictable, and brusque in its semantic effects, to that extent it is derivational; but there is no way to draw a consistent, motivated line such that all examples on one side are clearly and truly inflectional while all those on the other are clearly and truly derivational.

This has the very important implication that any theory that crucially depends on such a sharp distinction is operating on extremely shaky ground. The obvious examples are the many theories that hold that there is a sharp distinction between lexicon and syntax, and want to account for inflection in the syntax and derivation in the lexicon, or productive structures in syntax and non-productive structures in the lexicon, or predictable structures in syntax and non-predictable structures in the lexicon, with no overlap. The burden of proof is on such theories to show how a consistent line can be drawn, motivated by the nature of the distinction itself rather than the convenience of the theoretician. If this cannot be done, the idea of a sharp distinction must be abandoned as artifactual, and a more natural model substituted for it.

Notes

[1] This paper is an informal and undocumented exposition of how I have come to view an issue that is, to me, quite important. I do not claim to be the first to perceive these insights into the inflectional/derivational distinction: I do think that they need to be more widely apprehended by linguists.

[2] Where there is a plurality of parameters they tend to be naturally related. I.e., a high (or low) reading along one parameter will tend for some independent reason, relating to the nature of our cognitive system or of the experiences we process through it, to correlate with high or low readings along the other parameters. This is what permits the categories to be perceived as unified.

For instance, **big**, in the physical sense, involves size along the three dimensions of space and, to a lesser extent, the dimension of weight: a prototypically **big** man exceeds the norms for human size along all four parameters. They are naturally related in that the human body and many other objects to which we apply the term tend in our experience to preserve the same proportions and density at various sizes. Note however that a man might be **big** (in a less than prototypical sense) even if thin (as long as he was quite tall), or even if short (as long as he was very heavy and fat).

[3] This analysis is drawn from Ronald Langacker's Cognitive Grammar (née Space Grammar), of which Langacker 1982 and Langacker (in press) are perhaps the most accessible works.

[4] Such a difference would be interpreted by many linguists as a difference of syntactic function rather than a semantic variation, and therefore as evidence that we must posit two separate homophonous suffixes rather than a single suffix. (This position will be even easier to take with respect to later examples: e.g. to posit a different suffix **-er** in **locker** or in **footer** than in **swimmer**.)

My position (and that of Cognitive Grammar) is that there is certainly here a difference in the syntactic construction, which (as it must) means a difference in the semantic structure as well. It would be overly simplistic to ignore those differences, but it would be equally simplistic to therefore split the morpheme in two, ignoring the great similarities (semantic, syntactic, and phonological) uniting the two usages. A morpheme is not a monolithic structure, but it is unified and can be treated so for all that.

[5] The dictionary does give this as a possible meaning.

Worrisome can work either way in my dialect—a worrisome person is one who tends to worry too much; a worrisome development one that leads someone to worry about it. **Cuddlesome** means most saliently something like "tending to make one want to cuddle it", but most cuddlesome things are so partly because they tend to cuddle to one.

[6] Actually it might be to the verb **weary** rather than to the adjective. (In fact, on my view, it is probably correct to include both analyses in the grammar.) Similarly, **quarrelsome, burdensome, fearsome, frolicsome** and other forms may be derived from either the nominal or the verbal form of the stem, or both. (The same does not hold for forms like **cumbersome, tiresome, irksome,** and **cuddlesome**, which do not have the appropriate nominal forms.)

[7] This parameter of extent to which the meaning of the stem is changed might be viewed as in some sense primary over the others; it is the one the inflectional/derivational distinction is named for (inflection is "bending" the meaning of the stem, derivation is "deriving" a new, i.e. a different, stem). It is tied in naturally with the other parameters in a number of complex and subtle ways: e.g. a brusque change, as noted, encourages us to perceive the two forms as being different words/stems. This in turn encourages us to resist leveling tendencies and to permit semantic drifting, which then results in a loss of predictability.

It should be noted that, because of my theoretical bias, I am including the shifting of syntactic category of a form as a kind of meaning change; it is perhaps the prototypical kind of meaning change produced by derivational morphology. On this view, a change in the basic semantic organization of the stem (a shift in profile type, to use Langacker's term) is effected, which has natural consequences on the range of possible syntactic structures the complex stem can be used in.

[8] I do not want to claim that these are the only parameters relevant to the inflectional/derivational distinction. For instance, I have failed to mention the extent (often iconic for degree of semantic change) to which the morphology changes the phonological structure of the stem it attaches to, nor have I mentioned the tendency of agreement markings to be inflectional; I have also not discussed the loss of analyzability that tends to be more characteristic of derivational morphology than of inflectional (e.g. **handsome** or **drawer** are much less saliently analyzable than is **quarrels**). What I am claiming is that at least these three parameters are relevant, and that they are all scalar rather than dichotomous, and that

therefore the overall distinction is both complex and non-dichotomous.

[9] It is an oversimplification, of course, to represent the position of an entity as complex as **-er** by a dot: rather it occupies an area, with different *versions* of it (corresponding to different usages) being nearer to the inflectional or derivation pole than others, but all being unified as examples of the same morpheme. The dot we use in the diagram can be thought of as the "center of gravity" of the area covered by the morpheme in all its usages.

[10] It is certainly not irrelevant that this person is strongly affected by the amputation, and thus is indirect-object-like to some degree. Note also the phonological truncation of the stem from **amputate** to **amput-**; something similar is probably going on in the derivation of **conferee** from **conference.** I have heard the term **designee** used: this would be an even more direct parallel.

[11] In some cases it can be reasonably claimed that the nouns are first converted to adjectives by the (rather productive) rule involved in cases like **cotton shirt** or **blackboard chalk.** For instance, one (or at least I) could say "It´s sort of a **cottonish** material". In other cases (e.g. **childish, old-maid-ish**) such an analysis seems less likely; particularly this is so where the idea of attenuation of the quality designated is not involved. E.g., in **cottonish**, as in **reddish**, **-ish** means "having the quality of the adjective, but not strongly", whereas in **childish** or **old-maid-ish** it means "having a particular quality characteristic of the noun" whether strongly or not. This is why **very** collocates well with the second group but not the first: **very reddish** is somewhat odd, almost a contradiction in terms, but **very childish** is perfectly natural.

Another class of denominal adjectives in **-ish** which also must be included in a complete accounting is that of adjectives of national or cultural origin or character such as **Amish, Danish, English, Polish, Scottish, Spanish**, etc. This usage does not seem to be presently productive.

References

Langacker, Ronald W. 1982. Space grammar, analysability, and the English passive. Language 58.22-80.

———, in press. Foundations of cognitive grammar. Stanford University Press.

THE KOH VERBAL SYSTEM

Suellyn H. Glidden

1 Introduction
  1.1 Location and population of Koh
  1.2 Classification and earlier studies of Koh
  1.3 Basis and purpose of the present study
  1.4 Basic phonology
  1.5 Basic syntax
    1.5.1 Constituents and constituent ordering
      1.5.1.1 Clause level
      1.5.1.2 Noun phrases
    1.5.2 Clausal constructions
      1.5.2.1 Serial verbal constructions
      1.5.2.2 Multiple clauses
      1.5.2.3 Coordinate and subordinate clauses
2 Phonological and morphological structure of the verb
  2.1 Verb root morpheme structure
  2.2 Verbal affixes
3 Classification of verb roots
  3.1 Copular forms
  3.2 Transitivity
  3.3 Inherent aspect
4 Tense, aspect, and mood in Koh
  4.1 Inflection and mood distinctions
    4.1.1 Base tone and the factive mood
    4.1.2 High tone and the non-factive mood
      4.1.2.1 Stated possibilities
      4.1.2.2 Desire/volition/purpose
      4.1.2.3 Obligation
      4.1.2.4 Contrary-to-fact
      4.1.2.5 Hortative/imperative
    4.1.3 Rising tone
    4.1.4 Tone perturbation
    4.1.5 Summary: verbal inflection
  4.2 Auxiliaries and aspect distinctions
    4.2.1 **ddò** and ´PROGRESSIVE´
    4.2.2 **bba** and ´HABITUAL´
  4.3 Aspectual particles
    4.3.1 **bbày** and ´IMPERFECTIVE´
    4.3.2 **ró** and ´PERFECT´
5 Complex constructions: complement-taking verbs and serial verbal constructions
  5.1 Complement-taking verbs
    5.1.1 Equi-construction
    5.1.2 Manipulative with **bbày**
    5.1.3 Manipulative with **zàddká**
    5.1.4 Explicit benefactor and **ká**

- 5.2 Serial verbal constructions
  - 5.2.1 Linking SVCs
    - 5.2.1.1 **hą** ´to give--benefactive/causal´
    - 5.2.1.2 **vi̱** ´to come´
    - 5.2.1.3 **vǎ** ´to go´
  - 5.2.2 Modifying SVCs
    - 5.2.2.1 **òo** ´to be finished´
    - 5.2.2.2 **màh** ´to be sufficient´
    - 5.2.2.3 **ddah** ´to do´
  - 5.2.3 Summary
- 6 Summary of verbal forms and their primary meanings
  - 6.1 Terms and definitions of Koh verbal categories
    - 6.1.1 Inherent aspectual classes
    - 6.1.2 Mood categories
    - 6.1.3 Aspect categories
  - 6.2 Interaction of mood and aspect with states, events, and processes

1 Introduction

1.1 Location and population of Koh

The term Koh refers in this paper to a language and people located primarily in the cantons of Pao and Bouroum of the Baiboukoum Sous-Prefecture and the Logone Prefecture of Chad. There are also Koh speaking migrant communities in the vicinity of Garoua, Cameroon, and in the border area between Cameroon and Chad. Koh is spoken by an estimated 7-10,000 persons. On the basis of informal observation and partially completed dialect surveys one can distinguish two dialects of Koh, corresponding to the two cantons of Pao and Bouroum where the language is spoken.

The Koh within these two cantons tend to be quite isolated from one another and from the neighboring ethnic groups: the Ngambay and Sakpu in canton Bouroum and the Nzak-mbáí, Mbum, and Karang in canton Pao. The Koh may understand some of the language of the neighboring ethnic groups if their villages are close and there is contact, but, generally, monolingualism remains very high. Very few Koh understand or speak French and almost none understand Arabic, the two national languages of Chad.

The Koh are a patrilineal society and are animistic or Christian. They are almost exclusively agriculturists. The main crops are red or white millet and peanuts. Cotton is the cash crop, and for most, the sole means of gaining a cash income. Goats are kept primarily as symbols of wealth and as part of the bride price system. The Koh are facing change at an ever increasing rate. Outside contact is increasing and resulting in greater expectations for goods, medical aid, and education. This will lead inevitably to changes in the Koh culture and language.

1.2 Classification and earlier studies of Koh

Koh is to be identified as one of the "Lakka" languages, classified by Greenberg (1963) as a member of Group 6 in the Adamawa branch of the Adamawa-Eastern family of Niger-Congo. The members of Greenberg's Group 6 are Dama, Mono, Mbere, Mundang, Yasing, Mangbei, Mbum, Kpere, Lakka, and Dek. Samarin (1971:215-219) and Boyd (1974:15-19) give a more detailed subclassification of most of the Group 6 languages. Samarin excludes Mangbei and Yasing from this group and adds five languages from Central African Republic. He also notes the ambiguity of the term "Lakka," which is used pejoratively by Muslims to refer to the ethnic groups occupying the Adamawa plateau and its environs. The ethnic groups of this area use the term "Lakka" in reference to a dialect of the Chadic language Sara-Ngambay.

Boyd (1974:17,18) gives a more precise and thorough subclassification of Group 6 languages. He uses the term **Lakka** for the proto-language of a large number of Group 6 languages. Within Lakka he establishes two main subdivisions, Eastern and Western, which are further subdivided. His final classification is as follows:

A. Western
    1. Pan/Mono, Mundang
    2. Dama, Galke/Pormi, Kali

B. Eastern
    1. Mbum, Man, Mbere
    2. a. Pandjama, Touboro, Koh
        b. Ndó-Mbàli, Nzak-mbái, Pana, Ngoumi

Within the subgroup 2 of the eastern group of Lakka, Boyd establishes two subdivisions based upon two corresponding phonological sets:

Group a	**Group b**
w	v
v	g

Touboro is, as noted by Boyd, a dialect of Pandjama which is also known as Sakpu. Ndó-Mbàli is also known as Karang. Koh is classified in Group 2a, of the eastern division of Lakka, of Group 6 Adamawa languages, in the larger division Adamawa-Eastern of Niger-Congo.

There have been no previous linguistic studies of Koh aside from an unpublished comparative study of Koh and Karang by Ubels (1983). Ubels' purpose was to study the feasibility of establishing one standard orthography for both Koh and Karang. sixty-seven percent of the 487 words used to compare Koh and Karang are cognate. But while Karang is closely related to Koh, both phonologically and also syntactically, Ubels concluded that they were separate languages and would require separate writing conventions. Karang is the language most closely related to Koh in which there has been linguistic investigation (Ubels and Ubels 1980, Ubels 1983).

1.3 Basis and purpose of the present study

The data were collected between June 1983 and April 1984. I was aided in my investigation primarily by Bétar Ozias (canton Pao), Beassem Sylvain (canton Pao), Soroda Samuel (canton Bouroum), and Médangé Silas (canton Pao). These four men are French speakers and hold positions of leadership in their respective villages. They were chosen to be my informants with the approval of a committee of Koh representing both the cantons of Pao and Bouroum.

The purpose is to describe the Koh verbal system in terms of its grammatical forms and their possible meanings. Special attention is given to the tense-aspect-mood (**TAM**) system. The formally marked categories of the TAM system are relatively few in Koh. In fact Koh is a "tenseless" language in the sense that temporal deixis does not appear as a formal category in the verbal system. Rather, it is a sentence level function or is derived from the discourse-pragmatic context. There are one basic mood opposition and four aspectual categories. One aspectual category has a restricted usage. Even though there are a small number of overt forms in Koh, there is a wide range of expressible TAM notions. So, in order to describe the verbal system adequately, one must consider the full range of meanings possible for each form and combination of forms.

Before turning to the consideration of the verbal system I will introduce the basic phonology (Sect. 1.4) and syntax (Sect. 1.5) of Koh. Then in Sect. 2 I will examine the relevant phonological and morphological information concerning verb roots. In Sect. 3 the classification of verb roots will be discussed. Then in Sect. 4 I will turn to the formal TAM categories of Koh, and in Sect. 5 to complex verbal constructions. I will conclude, in Sect. 6, with a summary of verbal forms and some of their possible meanings.

1.4 Basic phonology[1]

There are thirty-five phonemes in Koh, of which seven are vowels and twenty-eight are consonants. These phonemes are represented in the orthography as follows.

i	voiced	high	close	front	unrounded	vowel
e	voiced	mid	close	front	unrounded	vowel
ee	voiced	mid	open	front	unrounded	vowel
a	voiced	low	open	mid	unrounded	vowel
u	voiced	high	close	back	rounded	vowel
o	voiced	mid	close	back	rounded	vowel
oo	voiced	low	close	back	rounded	vowel

b	voiced	bilabial		stop
bb	voiced	bilabial		stop-implosive
d	voiced	alveolar		stop
dd	voiced	alveolar		stop-implosive
f	voiceless	labiodental		fricative
g	voiced	velar		stop
g	voiced	labio-velar		stop
h	voiceless	glottal		fricative
k	voiceless	velar		stop
kp	voiceless	labio-velar		stop
l	voiced	alveolar		lateral
m	voiced	bilabial		nasal
mb	voiced	bilabial	prenasalized	stop
mgb	voiced	labio-velar	prenasalized	stop
n	voiced	alveolar		nasal
nd	voiced	alveolar	prenasalized	stop
nz	voiced	alveolar	prenasalized	fricative
ŋ	voiced	velar		nasal
ŋg	voiced	velar	prenasalized	stop
p	voiceless	bilabial		stop
r	voiced	alveolar		retroflex
s	voiceless	alveolar		fricative
t	voiceless	alveolar		stop
v	voiced	labiodental		fricative
vb	voiced	labiodental		flap
w	voiced	labio-velar		non-syllabic vocoid
y	voiced	palatal		non-syllabic vocoid
z	voiced	alveolar		fricative

All examples of Koh data in this paper are written according to the proposed standard orthography.

In addition to these phonemes, Koh has three level tones: high, mid, and low, which are symbolized as ´, unmarked, and ` respectively. There is also a rising contour tone, symbolized as ˇ. There is one phonemic tone per morpheme. Tone is marked on the first vowel of each morpheme.

There is nasalization prosody in Koh which applies to the word. In every word where nasalization occurs, the nasalization extends to every vowel and to certain consonants. For example, in **wa̧ra** ´man´, both vowels and **w** and **r** are nasalized. Nasalization is symbolized by a cedilla (,) under the first vowel of the word.

1.5 Basic syntax

Before I turn to the description of the verbal system, I will briefly outline in this section the main clause and noun phrase constituents and their ordering, and several clause level constructions.

1.5.1 Constituents and constituent ordering

1.5.1.1 Clause level. Koh is an SVO language.[2] The basic constituent order in intransitive sentences is SV, as seen in (1), and in transitive sentences is SVO, as seen in (2).

(1) mì zool
I go
´I´m going.´

(2) mì nza mbedde
I look paper
´I´m looking for paper.´

Indirect objects and other obliques usually occur after the direct object or after the verb when there is no DO (3).

(3) mì se bbà Laramanay
I go to Laramanay
´I´m going to Laramanay.´

Other main clause constituents usually occur at the end of the clause. The basic order of these constituents is: adverbs, aspectual particles, negative marker, and question marker. While it is unusual to have all of these constituents simultaneously, (4) is a maximal expansion showing the constituent ordering.

(4) i bbàh bbaỳ bbamba bbấy ya le
they talk word a:lot IMPFT NEG QM
´They aren´t still talking a lot, are they?´

Emphasis by left-dislocation is frequent in Koh for object NPs, temporal adverbs, or prepositional phrases.

(5) fè kếy nấ, mì su̧ ya
thing this here I eat NEG
´I don´t eat that.´

1.5.1.2 Noun phrases. Modifiers typically follow the noun which they modify. Within the noun phrase there is a general fixed constituent order as follows: NOUN (ADJ) (POSS) (PL) (NUMBER) (REL CL) (DEM/LOC). The number of modifying constituents within a noun phrase rarely exceeds three with, possibly, a relative clause as a fourth (e.g., (6) as a maximal expansion NP).

(6) nzòbb kpɨ́ri-rì síddi ká i guh fè
person leprous-PL two REL they groan thing
'the two lepers who are groaning'

Koh has both alienable and inalienable possession, reflected in the following two sets of possessive pronouns:

Alienable		**Inalienable**
bbi	my	-i
bbo	your	-a
bbe	his	-e/ee
nári	ours (incl)	nári
ná	ours (2)	ná
bburu	ours (excl)	bburu
bbarì	yours (PL)	bbarì
bbari	theirs	bbari

1.5.2 Clausal constructions.

1.5.2.1 Serial verbal constructions. The juxtaposition of two or more verbs with a common subject in a single clause occurs in Koh, as it does in a large number of Niger-Congo languages. Welmers (1973:366) calls this phenomenon "serialization." The structure of serial verbal constructions is discussed in Sect. 5.2.

1.5.2.2 Multiple clauses. States or actions associated with a single subject which are not expressed through a serial verbal construction are frequently expressed in multiple clause constructions. In multiple clauses the first clause contains reference to the subject, and all subsequent clauses in the construction have a corresponding subject pronoun. The form of subject pronouns in multiple clauses differs from their form in independent clauses, as shown in the following chart:

Subject Pronouns

Independent Clause		Multiple Clause
mì	I	mí
mù	you	mú
ka	he	à
ną́ri	we (incl)	ną́ri
ná	we (2)	ná
bburu	we (excl)	bbúru
ì	you (PL)	í
i	they	í

Actions expressed in multiple clauses are either simultaneous or sequential. Example (7) illustrates a multiple clause construction where the actions are simultaneous. In (8) the actions are sequential.

(7) i mąy bbày, í mbi nzáh-rì, í
they discuss word they take tongue-PL they

bbàh míh
talk that

´They discussed, they had their say, and they said....´

(8) ka mgba bàl ná, à i ni, à gòodd ni
he seize goat there he kill it he gut it
´He grabbed the goat, killed it, and gutted it.´

1.5.2.3 Coordinate and subordinate clauses. Coordinate and subordinate clausal constructions are a significant part of Koh syntax. Clauses in conversation or discourse seldom occur without a conjunction of coordination or subordination which relates what is said to previous and/or following statements. Subordinate clauses generally occur before the main clause. Subordination markers occur clause initially (9).

(9) zàddká mì bba lèdd hóy ná, bǎh bbáy máh,
when I DCOP child only there father and mother

i fère mì bbày ŋgéereewų́ru
they teach me word God

´When I was just a child, my parents taught me the Bible.´

Coordinating conjunctions may occur sentence initially, relating what follows to a previous statement, or between two clauses forming a continuum of thought or action. In example (10) **báyɨh lee** ´and then´ relates the clause to the previous clause (which in this situation was a direct quote), and **lee** ´and´ joins clauses that form a continuum of thought. This sentence is from a narrative text.

(10) báyɨh lee ná tạwara ví mgba seri mbel, à
then and there panther come seize tooth pig he

kùoo lee nah wukluk, à lieke tieeŋsee lee
pull and ?it come:out he fall backwards and

sàw-e ví haw dduoomì
tail-his come break silently

´Then Panther came and grabbed Pig´s tooth; he pulled on it and it came out; he fell backwards and his tail broke silently.´

Complement subordinate clauses are considered further in Sect. 5.1.

2 Phonological and morphological structure of the verb

2.1 Verb root morpheme structure

There are six consonant/vowel patterns in Koh verb roots: (C)V, (C)V_1V_2, (C)VC, (C)VCV, and CVCCV. The majority of Koh verb roots are monosyllabic. Each verb root, whether monosyllabic or disyllabic, has one lexical tone. This base tone is either mid or low. The base tone of a verb is most clearly seen in a negated factive sentence.[3]

Reduplication of the verb is common in Koh. The initial (C)V of the verb root is reduplicated and preposed to the root. The lexical tone of the verb is carried by the reduplicated (C)V and the root´s tone is neutralized to mid. This process is illustrated in (11) and (12) below, where the unreduplicated verb root in (11) is reduplicated in (12).

(11) tédd hùo bbamba
cold:N cold:V very
´It is very cold.´

(12) tédd hù-huo
cold:N RED-cold:V
´It is cold.´

The entire verb root is reduplicated when an object is present (13). This object, whether noun, noun phrase, or pronoun, occurs between the reduplicated particle and the verb root.

(13) tédd hùo mì huo
cold:N RED I cold:V
'I'm cold.'

2.2 Verbal affixes

There are two verbal affixes, **-ka** and **-ke**. **-ka** occurs with certain CV verb roots, but only when the verb carries a high tone. When **-ka** is affixed to a verb root, the resultant form is a variant of the verb root.

In Karang, a similar particle, **-na**, occurs and its function is to nominalize intransitive verbs.[4] That **-ka** cannot have the same function in Koh is apparent for two reasons. First, a Koh verb with **-ka** can appear as a main verb in a clause (14) and (15).[5]

(14) mì sų́ŋa fè
(sų́-ka)
I eat thing
'I'm going to eat something.'

(15) máh, à kɨŋa mbedde bbi
(kɨ-ka)
mother she read paper my
'My mother will read my letter.'

When **-ka** occurs with a verb root, it does not have an identifiable meaning and does not modify the meaning of the verb root. In (15) **-ka** was affixed to the verb **ki** 'to read'. It is possible to omit **-ka** without changing the meaning (16).

(16) máh, à kɨ mbedde bbi
mother she read paper my
'My mother will read my letter.'

A second reason for distinguishing the function of **-ka** in Koh from that of **-na** in Karang is that the affixation of **-ka** to verb roots is not restricted to intransitive verb forms. It appears, for example, with the locative copula **ddò** 'to be located', the descriptive copula **bba** 'to be', and is obligatory with such transitive forms as **mbi** 'to take', **mgba** 'to hold', and **koo** 'to see or know'.

All verb roots taking the affix **-ka**, whether optionally or obligatorily, when they carry a high tone are short open CV syllables. The only exception to this is the verb **yạh** 'to take', for which no contrastive short form occurs. The affix **-ka** is obligatory with **yạh** when it has a high tone. Aside from **yạh**, no non-CV verb root occurs with **-ka**.

(17) ka séke kpʉru bbày yⱥ́hŋa (yⱥ́h-ka) zàdd bbe
he go until that take place his
'He will advance until he takes his place.'

-ka usually, but not always, appears in the nominalized form of the verb roots with which it occurs, as the following examples show:

yⱥh	'to take'	yⱥ́hŋa kɨ bburu REFL our	'our marriage'
vbu	'to throw'	vbúka	'the throw'
vi	'to come'	sie víka hour	'the arrival
kɨ	'to read'	fè kɨ́ŋa thing	'reading material'

The only other affix that occurs on the verb is a complement marker, **-ke**, which indicates an underlying oblique object that does not appear on the surface. The verbal complement usually is one of instrument, 'by means of'. Example (18) illustrates **-ke** to mark an unspecified oblique, (19) makes explicit the oblique, and (20) shows the impossibility of **-ke** co-occurring with the oblique phrase.

(18) mì húŋn-ke gári
I rub-OBLM clothes
'I will rub the clothes (with something).'

(19) mì húŋn gári bbáy sàbun
I rub clothes with soap
'I will rub the clothes with soap.'

(20) *mì húŋ-ke gári bbáy sàbun
I rub-OBLM clothes with soap
'I will rub the clothes with soap.'

3 Classification of verb roots

Koh verb roots may be classified by the traditional categories of copula, intransitive, and transitive; see Sect. 3.1 and 3.2. Further, as discussed in Sect. 3.3, inherent aspectual differences in verbs must be carefully studied in order to adequately describe TAM categories and verbal systems.

3.1 Copular forms

There are two copular forms in Koh, each with several distinctive functions. The copula **bba** 'to be' (glossed DCOP) is used in two constructions to express identification and equation/characterization. A copular sentence of identification is formed by the use of **bba** followed by a nominal complement.

(21) bba lédd
DCOP child
´This is a child.´

An equation/characterization sentence is formed by a nominal phrase followed by **bba** and a nominal or adjectival complement.

(22) vúri bba nám pédd mbɨ́w
today DCOP day work one
´Today is Monday.´

(23) mì bba gèl
I DCOP left
´I´m left-handed.´

Koh also has a locative copula **ddò** (glossed LOC) used to indicate location. This copula can be followed by a nominal complement (24) or an adverbial complement (25).

(24) máy bbi ddò hùl-a
wife my LOC house-her
´My wife is at her house.´

(25) máy bbi ddò ddi
wife my LOC there
´My wife is there.´

As in many Niger-Congo languages (Welmers 1973:309), the locative copula can be used to express possession.

(26) mì ddò bbáy lédd-rì
I LOC with child-PL
´I have children.´

3.2 Transitivity

Few Koh verbs are intransitive. These can be subdivided into three groups. The majority are stative and must be translated into English by ´to be´ and some attribute. **hìh** ´to be heavy´ in (27) and **ndàdd** ´to be good´ in (28) are examples of stative verbs that are intransitive.

(27) hìh bbamba
heavy very
´It is very heavy.´

(28) ngáw bbo ná ndàdd bbamba
food your there good a:lot
´Your food is very good.´

These forms are full-fledged verbs: they accept modal and aspectual categories and can be reduplicated or negated as any other verb. Example (29) illustrates a reduplicated verb, while (30) illustrates negation.

(29) ngǻw bbo ná ndà-ndadd
food your there RED-good
´Your food is good.´

(30) ngǻw bbo ná ndȁdd ya
food your there good NEG
´Your food isn´t good.´

These verbs are the only verbs, other than certain copular constructions as in (21), for which a subject is not obligatory (27).

Verbs of movement form a distinct semantic class in Koh. Only two have been found that never take an object (**zool** ´to go or leave´ and **na** ´to remain´). All others are more correctly called semi-transitive in that they optionally take either a direct object (31a) or an oblique (31b).

(31) a. mì se Laramanay
I go Laramanay
´I´m going to Laramanay.´

b. mì se bbà Laramanay
I go to Laramanay
´I´m going to Laramanay.´

A third subclass of intransitive verbs, a closed class, includes verbs describing the weather and verbs in which the subject is always a patient and not an agent.

(32) mbàm tȍo-too
rain:N RED-rain:V
´It´s raining.´

(33) pédd bbi ȍo ró
work my done PERF
´My work is already done.´

The majority of Koh verbs are transitive, requiring a direct object.[6] Most transitive verbs are expected semantically to have a direct object. Such verbs as **siŋ** ´to attach´, **gaŋ** ´to wait for´, **loŋ** ´to bite´, and **nza** ´to look for´ are representative of this open class. More interestingly, there is a set of transitive verbs in Koh that, semantically, would not normally be expected to require a direct object. This closed subclass of transitive verbs has what Welmers (1973:484) calls a "cognate object." This cognate object is a nominal which is derived from the verb form for which it is the object. This nominal appears only as the object of the verb form from which it is derived. Cognate objects can be distinguished from a reduplicated verb root by their tone. When a verb is reduplicated the reduplicated portion carries the tone and the root is neutralized to mid. A cognate object carries its own lexical tone. Examples of verbs with cognate objects are as follows:

bbàh bbày	´to speak a word´
ddù ddúl	´to run a run (race)´
na nám	´to sleep a sleep´
pah pày	´to cultivate a cultivation´
ree rèew	´to cry a cry (tears)´
se sédd	´to walk a walk´

Further, many verbs with a cognate object can also be reduplicated, as in (34) and (35).

(34) mì bbàh bbày bbah
I RED word talk
´I´m talking.´

(35) mì pah pày pah
I RED field plant/cultivate
´I´m cultivating.´

There are methods by which a direct (including a cognate) object may be deleted from an ordinarily transitive sentence. The most common use of reduplication is as follows:

NP V NP -> 1 Redup + 2 ∅
1 2 3

Examples (36a) and (36b) illustrate this process.

(36) a. mì su̩ fè
I eat thing
´I´m eating something.´

b. mì su̩-su̩
I RED-eat
´I´m eating.´

A grammatical function of reduplication is to allow for the deletion of an object in a normally transitive sentence and thus reduce its valency.[7] There is also a pragmatic function: the reduplication may give assertive focus to the action specified by the verb. The object may also be omitted in the presence of the following verbal or clausal elements: adverbs (37b), aspectual particles (37c), negation (37d), or interrogation (37e).

(37) a. mì gaɲ nzòbb
I wait person
´I´m waiting for someone.´

b. mì gaɲ hőy
I wait only
´I´m only waiting.´

c. mì gaɲ bbấy
I wait IMPFT
´I´m still waiting.´

d. mì gaɲ ya
I wait NEG
´I´m not waiting.´

e. mì gaɲ le
I wait QM
´I´m waiting?´

In all these cases, there appears to be a special focus on the action of the verb which overrides the normal presence of an object.

3.3 Inherent aspect

The distinction copula/non-copula and transitive/intransitive are useful natural classes; I have shown some of their distinctive characteristics above. However, of greater significance for the classification of Koh verbs is their inherent aspectual meaning. It is necessary to consider these; for when inherent aspect interacts with the syntactic TAM categories, especially the aspectual categories, the meaning of the syntactic category and/or the inherent aspect of the verb may change. Thus, the meaning of an aspectual marker may vary according to the inherent meaning of the verbs with which it is combined. The possible meaning of a verb may also change as it combines with different aspects. The interaction of inherent and syntactic aspect may also prohibit or restrict the co-occurrence of certain forms.

For the purpose of this paper, aspect is defined as the manner of "viewing the internal temporal constituency of a situation" (Comrie 1976:3). Syntactically, aspects are limited to the number of grammatical forms which express them. It is more difficult to form a concrete list of inherent verbal aspects since semantic notions vary widely. Lyons (1977:706) defines the inherent aspect of a verb as, "that part of its meaning whereby it (normally) denotes one kind of a situation rather than another." He goes on to say that many languages differentiate among states, events, and processes. These distinctions seem to reflect natural classes in Koh and are useful in analyzing inherent and grammatical aspect. Events and processes denote dynamic situations, whereas states refer to non-dynamic situations.[8] Events and processes are distinguished in terms of perfectivity. Typically, verbs that make events view the situation perfectively, whether or not the action expressed by the verb is punctual or durative with a built-in termination point. Processes, on the other hand, view the situation imperfectively. States, events, and processes are the three inherent aspectual classes with which I deal in this paper.[9] Comrie (1976:13) summarizes these classes as follows: "States are static, i.e. continue as before unless changed, whereas events and processes are dynamic, i.e. require a continual input of energy if they are not to come to an end; events are dynamic situations viewed as a complete whole (perfectively), whereas processes are dynamic situations viewed in progress, from within (imperfectively)."

States differ from events or processes in that a stative situation is non-dynamic. Inherently stative verbs, then, would refer to durative situations that lack any element of progressive change. Examples of verbs with a primarily stative meaning in Koh are seen in the following examples:

(38) mì koo fè
I know thing
'I know something.'[10]

(39) mì hɨ̀-hɨee
I RED-fear
'I'm afraid.'

(40) mì ker bbày
I think word
'I'm thinking.'

(41) sò-i sée mì see
body-my RED me hurt
'I'm sick.'

Events refer to dynamic situations that are viewed as a whole. There may be several subcategories within events, but the primary focus in this study is on **telic** situations in which there is a built-in terminal point beyond which the action cannot be said to continue. Representative verbs in this class are seen in these examples:

(42) ka mboŋ lédd
she birth child
'She had a child.'

(43) ka mbuo hùl
he build house
'He built a house.'

(44) mì zool
I go
´I´m leaving.´

Processes differ from events in that, while they refer to dynamic situations, they have no intrinsic terminal point. They are atelic in that the action, which is inherently durative, could conceivably continue indefinitely or it could end at any point and an instance of that action would have occurred. All verbs in Koh taking cognate objects are process verbs (see Sect. 3.2). Other examples in Koh are as follows:

(45) ka ru sùnum
he plant peanuts
´He is planting peanuts.´

(46) ka ddaŋ mbìh
he draw water
He is drawing water.´

(47) ka ddi sìm
he sing song
´He is singing.´

My interest in these inherent aspect classes in Koh is to study them in relation to syntactic TAM categories, some of which are naturally compatible (process verbs with the progressive), others incompatible (stative with the progressive or the imperfective). In order to understand adequately the range and semantic meaning of grammatical forms in Koh, they must be studied with constant reference to the inherent meaning of the verbs with which they appear.

4 Tense, aspect, and mood in Koh

Tense, aspect, and mood (TAM) distinctions are made by means of verb inflection, auxiliaries, and aspectual particles. In the following sections I examine the grammatical forms used and the range of meanings associated with these forms. First, I consider verbal inflection (Sect. 4.1), which expresses mood distinctions. Within this section I also examine two complications in verbal inflection: an unusual rising tone (Sect. 4.1.3) and tone perturbation (Sect. 4.1.4). Two auxiliaries are described in Sect. 4.2 which express aspectual categories. Aspectual distinctions are also expressed by aspect particles. These particles and their meaning are studied in Sect. 4.3.

4.1 Inflection and mood distinctions

Koh verbal inflection is limited to tone, which expresses mood. Mood here refers to the grammaticalization of the relationship between the speaker and his message (i.e., whether the speaker is affirming, commanding, asking) or of the relationship of the message´s content to reality (i.e., true, possible, probable, necessary). Thus, grammatical categories that distinguish degrees of actuality or desirability of an action are mood categories.

Tone is used in Koh to make one basic mood distinction, which Lyons (1977:816) refers to as the distinction between factive (F) and non-factive (NONF). This distinction has been noted in other Niger-Congo languages, and various terms have been used to express the same phenomenon: real/non-actual, actualized/non-actualized, realis/irrealis, etc. I have chosen to use the terms factive and non-factive.

All Koh verbal roots fall into one of two classes according to their tone in their most unmarked form. This **base** tone is either mid (**koo** in (48)) or low (**ddàŋ** in (49)) and marks factivity. Factivity expresses the actuality of a situation at the time of utterance or time of reference made clear in the utterance.

(48) mì koo fè
I know:F thing
´I know something.´

(49) mì ddàŋ mbìh
I draw water
´I draw water.´

Non-factivity is indicated by high tone for all verbs. Non-factivity asserts a meaning of potentiality of a situation.

(50) mì kóo fè
I know:NONF thing
´I´ll know something.´

(51) mì ddáŋ mbìh
I draw:NONF water
´I´ll draw water.´

These two tone classes, base tone (mid or low) and high tone, express the basic mood distinction between factive and non-factive. The purpose of the following discussion is to further show the range of semantic functions of these two basic mood categories.

4.1.1 Base tone and the factive mood. The factive mood marks a situation that is either fully developed at the time of utterance or has reached a point of stable development at the moment of speaking (Johnson 1980:227). While this is basically a mood distinction, it also implies reference to past or present time. Depending on the inherent aspect of a verb, factive mood indicates past or present, as now explained.

An inherently stative verb in the factive mood with no marked aspect will usually be interpreted, if the context allows, as referring to a situation that is a present reality. For example, **hɨee** ´to fear´ is stative and is interpreted as a present condition when combined with factive mood in (52).

(52) mì hɨ̀ee fè hɨee
I RED thing fear:F
´I´m afraid of it.´

In the case of an event verb which is inherently telic, however, factivity is generally interpreted as indicating an accomplished fact, and is thus translated with an English past tense (53).

(53) ka mboŋ lédd
she birth:F child
´She had a baby.´

Process verbs which center in the internal structure of a dynamic situation are more difficult to interpret as either past or present reality. In the absence of any overriding factor (aspect markers or time reference), an inherent process verb in the factive mood may refer to either a past or present situation and either a completed or a continuous action. Thus, (54) can be interpreted variously as a) past continuous, b) past action, c) present continuous, d) present action.

(54) ka ru sùnum
he plant:F peanuts
´He was planting peanuts´, ´He planted peanuts´,
´He is planting peanuts´, or ´He plants peanuts
(general truth).´

Although these are the most general interpretations associated with base tone on stative, event, or process verbs, factivity can express other meanings.

Factive mood on a stative verb may also refer to a fully developed state in the past (55) or even to a general truth (56).

(55) liee, mì hɨ̀-hɨee
yesterday I RED-fear:F
´Yesterday, I was afraid.´

(56) mì hɨ̀ee ndáy-rì hɨee
I RED cow-PL fear:F
´I fear cattle.´

Event verbs which refer to dynamic telic situations may also be interpreted as a present point of action. Thus in answer to the question, "What is he doing?" one can respond:

(57) ka mbuo hùl
he build:F house
´He´s building a house.´

It is also possible to interpret factive mood with an event verb as habitual or general truth, although in this case the inherent perfectiveness of the verb is lost. Thus one can ask, "What do the Koh do in January?" and get the response:

(58) i mbuo hùl
they build:F house
´They build houses.´

As indicated above, inherent process verbs can also have this habitual or general truth interpretation in the factive mood. The form of (54) above could be given in reply to the questions, "What does he do in May?" or, "What is his job?"

There is a case in which a verb marked for factive mood (i.e., with the base tone) is used even though the indicated action has not actually occurred at the time of speaking; seemingly, the speaker is so sure that the action will take place that he speaks of it as an actuality. For example, in (59) the speaker is asserting the certainty of his imminent departure by using the factive mood for an event that is not yet a fact. The equivalent in English might be, "I´m already gone."

(59) mì zòol ró
I go:F PERF
´I´m leaving.´

Base tone (mid or low) on Koh verbal roots is the grammaticalization of the factive mood, which basically says, "It is a fact that. . . ." Determining the exact meaning of the verb from among the wide range of possibilities is dependent upon the inherent meaning of the verb, the context in which the verb is employed, and the presence or absence of grammatical aspect markers. A discussion of concrete aspect markers and their interpretation with mood categories is the subject of Sect. 4.2 and Sect. 4.3. In the absence of marked aspect or limitations of context, factive mood can be interpreted as follows:

1. Present continuous
2. Present
3. Past continuous
4. Past
5. Habitual
6. General truth

Which of these various interpretations is the most natural rendering depends on the inherent meaning of the verb. Inherently stative verbs in the factive are usually interpreted as a stable situation in the present. Telic verbs are interpreted as fully developed action at the time of speaking, and, thus, are past events. Verbs which center attention on the inner dynamics of an action or situation are usually interpreted as either past or present continuous, the choice between these being determined either by contextual or overt markers of time and/or aspect. In summary, factive mood limits the action to the real world of **fact**, with various interpretations possible in that world.

The complement of factive is the non-factive mood (NONF). I now present the way in which non-factive is formally marked and the range of semantic interpretations of the verb in the non-factive mood.

4.1.2 High tone and the non-factive mood. Non-factive mood is indicated by high tone which is placed on the main verb of a simple clause or on the main verb of a subordinate clause in a complex sentence.

Non-factivity is used in Koh to assert possibility or potentiality. Verbs indicating situations which may be further developed are often interpreted as future. One of the functions of non-factive mood is to assert the likelihood of an action occurring in the future. Non-factivity, however, should not be identified with simple future tense: a simple assertion that something is likely to occur reflects the attitude of the speaker concerning the action and is thus a distinction of mood. In addition to stating possibilities, non-factivity is used to express desire/volition/purpose, obligation, contrary-to-fact conditions, as well as hortatives and imperatives. I will now examine briefly each of these semantic distinctions in the use of non-factive mood.

4.1.2.1 Stated possibilities. A simple clause with high tone on the verb is usually interpreted as a possibility (with various degrees of likelihood) unless the context indicates a different interpretation. This is true whether the speaker is asserting the potential for further development of a state (60), an event (61), or a process (62).

(60) mì hɨ́ee fè hɨee
I RED thing fear:NONF
'I will be afraid' or 'I may be afraid.'

(61) mì mbóŋ lédd
I birth:NONF child
'I will have a child.'

(62) ka ddáŋ mbìh
he draw:NONF water
´He will draw water´ or ´He may draw water.´

Interestingly, in Karang the various meanings of non-factive mood can in many cases be differentiated syntactically (Ubels 1983:12-26). Stated possibilities are marked by high tone on the verb, whereas in the case of other meanings the high tone is marked within the subject nominal phrase. Ubels classifies stated possibilities (called "non-subjunctive") as predictive and non-predictive. Predictive indicates not only the possibility but the likelihood of an event and is marked by a pre-verbal element **bbày**. This distinction is not found in Koh. Ubels (1983:13-16) groups other uses of non-factive mood under the term "subjunctive" because of the consistent manner in which the high tone is marked in the subject phrase. These syntactic differences do not appear in Koh. Because semantic distinctions are not formally marked in Koh as they are in Karang, I have chosen not to further subclassify the non-factive mood.

4.1.2.2 Desire/volition/purpose. Desire and purpose are expressed in Koh through complex sentences in which non-factive high tone is marked on the verb in a complement clause. Desire is expressed with complement-taking verbs such as the following:

Object Complement-taking Verbs

hìh bbày	´want that´
vbi bbày	´ask that´
gooŋŋ bbày	´beg, demand that´

Non-object Complement-taking Verbs

màh bbày	´be sufficient that´
ndàdd bbamba bbày	´be very good that´
ddò bbày	´be good that´
ddò ŋgooŋ bbày	´be important that´

These complement-taking verbs and their sentence structures are treated in greater depth in Sect. 5.1. The following examples show the use of the high tone on the subordinate verb to mark an unaccomplished desired action.

In object complement-taking verbs it is the desire of the subject of the main clause that is being expressed. The subject may desire to perform an action himself (63) or desire another to perform it (64). In both instances, the main verb of the subordinate clause is non-factive.

(63) ka hìh bbày séke zìŋ mù
he want:F that go:NONF with you
´He wants to go with you.´

(64) ka hìh bbày hą́ mù kaw
he want that give:NONF you stay
´He wants you to stay.´

In non-object complement-taking verbs it is the opinion or desire of the speaker that is recorded. Sentences with these types of verbs usually follow the pattern of (65).

(65) ndàdd bbày hą́ mù kaw
good:F that give:NONF you stay
´It´s good that you stay.´

As seen in the examples above, factive mood, marked by base tone on the complement-taking verb, represents an action/state that has reached a stable point at the time of speaking. The non-factive mood of the second verb marks a situation which has further potential for development and which, in this case, is desired. It is also possible to place high tone on the complement-taking verb. This results in a not-yet-accomplished state of desire, as shown in (66), where both the desire and the desired action are unrealized at the moment of speech.

(66) mì híh bbày séke zìŋ ni ròobbaáy
I want:NONF that go:NONF with him again
´I will want to go with him again.´

The complements of sentences expressing purpose have the same structure as complements of complement-taking verbs. But the matrix clause has a subject performing an action designed to result in the desired action/state indicated by the subordinate clause. Thus, in (67) the purpose of the subject´s going is to cultivate. In (68) the subject does something in order to cause someone else to do something.

(67) ka se bbà wáh bbày páh pày
he go:F to field that cultivate:NONF cultivation
´He went to the field in order to cultivate.´

(68) mì tą bbày háŋa rì hą mì nàŋ
I stop:F that give:NONF you:PL give:F me millet
´I´m stopping so that you can give me some millet.´

In both complements of complement-taking verbs and purpose clauses, high tone on the subordinate verb is used to mark non-factive mood.

4.1.2.3 Obligation. Non-factive mood can also be used to indicate that the referent of the subject is obliged to perform the action of the verb. In this case, it is due to some unnamed circumstance or the will of a person. In (69) the non-factive high tone on **seke** ´to go´ indicates the obligation placed on the referent by the will of the chief, **mbáy.**

(69) zàddká mbáy ddií mù lee, mù séke
when chief call:NONF you CONJ you go:NONF
´If (when) the chief calls you, you must go.´

In (70) the obligation is due to unnamed circumstances.

(70) mì hɨ́ee ŋgéereewʉ̀ru hɨee
I RED God fear:NONF
´I must fear God.´

4.1.2.4 Contrary-to-fact. Non-factive mood can be used to indicate contrafactual meaning in the past. In a simple clause the main verb carries a high tone.

(71) ka kér bbày
he think:NONF word
´He should have thought.´

In contrary-to-fact conditions the high tone is on the main verb of the matrix clause (e.g., **léke** in (72)).

(72) zàddká à se púoo ná ya lee,
if he go:F home there NEG CONJ

sòoy, à ti léke ni ya
snake he NEG bite:NONF him NEG

´If he had not gone home, the snake wouldn´t have bitten him.´

4.1.2.5 Hortative/imperative. A speaker may express his desire that something occur by a direct statement with the non-factive mood. Non-factivity indicates that the situation is outside the control of the speaker. Example (73) is such a hortative clause.

(73) ŋgéereewʉ́ru, ka sádd zìŋ mù
God he be:NONF with you
´May God be with you.´

A final use of the non-factive mood is in imperative clauses. Imperatives have been analyzed here as non-factive, although their forms do not always follow the pattern of high tone on the verb, as seen in other non-factive forms. I now present the form of imperatives and then the reasons for considering them non-factive.

Imperative forms are of various kinds depending on the base tone of the verbs. Those with low base tone form two classes. Those with mid form a third. The majority of low tone verbs form the imperative by a) raising the inherent low tone of the second person pronominal subject to mid and b) marking the verb with a high tone. Examples (74a) and (74b) show the formation of the imperative for Class 1 low verbs.

(74) a. mù rùma hùh
you put:out:F fire
'You are putting out the fire.'

b. mu rúma hùh
you put:out:NONF fire
'Put out the fire!'

A smaller class of low tone verbs (a closed class) form the imperative by raising the tone of the pronoun to high without modifying the tone of the verb. Class 2 low imperative formation is illustrated in (75a) and (75b).

(75) a. mù yɨ̀h
you come:F
'You returned.'

b. mú yɨ̀h
you come
'Come back!'

Mid base tone verbs form the imperative by raising the tone of the pronoun to mid with no modification of the tone of the verb. Mid base tone imperative formation is illustrated in (76a) and (76b).

(76) a. mù bber túl tagra
you nail:F head table
'You are nailing the table together.'

b. mu bber túl tagra
you nail head table
'Nail the table together!'

In each of the above rules there has been a raising of tone; this suggests a relationship with non-factive marking, which always involves tone raising. Class 1 low is typically non-factive with the high tone marked on the verb, while Class 2 low includes high tone, although it is on the pronoun. It is only mid base tone verbs that are ambiguous, since the pronoun is raised only to mid and not to high and the tone of the verb is not changed at all. In any case, the imperative forms definitely do not follow a normal factive form, but rather are more like non-factive.

There is an exact correspondence between Class 2 low[12] imperatives in Koh and all imperatives in Karang. Imperative forms in Karang are unambiguously non-factive. This supports the conclusion that Koh imperative forms should be interpreted as non-factive.

Finally, imperative as a notion fits into a non-factive mood category. We have described non-factive as referring to situations in which there is further potential for dynamic development. Non-factive can also express the speaker´s desire that an action occur. Non-factive can attest a lack of control of the speaker over the situation. All three of these criteria apply to an imperative clause. In a mood system such as Koh, where there is clearly a distinction between factive and non-factive, imperatives are non-factive.

4.1.3 Rising tone. We stated in the previous section that base tone (mid or low) and high tone are the inflectional forms making the mood distinction between factive and non-factive. I have occasionally noted (when eliciting text material) a rising tone on the verb when there seem to be semantic elements of both factivity and non-factivity. This rising tone appears only infrequently, and I have not as yet been able to predict its occurrence or find a frame in which it could be elicited. A rising tone appears in factive situations in which the action has started. The situations are also non-factive in that the actor expects or intends that this action will continue. Rising tone thus indicates present actuality with future intention. Several examples of rising tone are given here to show how factivity and non-factivity are both semantically present on one verb. In (77) **kàw** ´to stay´ has a rising tone.

(77) toom rǒo bba hỵ̀ru-sùo-ná roo zé ná
cotton first CDOP force-body-our ? TOP we

kǎw kéy yah
stay here !

´It´s the cotton that´s our strength that we can stay here!´

In (78) **bbòodd** ´to beat´ is marked with a rising tone.

(78)

Súh	ná	ví	mbi	pùh,	à	bbǒoddke	ni
Suh	there	come:NONF	take:F	wood	he	beat	him

kpőy-kpőy,	à	i	ni
pow-pow	he	kill:F	him

´Suh there gathers up a stick and beats him with it "pow-pow," and he kills him.´

In (79) **yą̀h** ´to take´ has a rising tone.

(79)

bbày	sù-oo	seereem	bbe	zé	hą	ni	yą̌hke
because	body-his	hurt	his	TOP	give:F	him	take

są̀w-pùh
root-wood

´Because of his sickness, he is taking medicine.´

In all phrases with rising tone on the verb, this rising tone may be replaced by the base tone. If so, the desirability of continuing the action is de-emphasized, and the verb simply describes the present situation.

Although further work is needed to isolate the factors controlling the use of rising tone, my preliminary analysis is that rising tone combines the semantic notions of factivity and non-factivity on one verb.

4.1.4 Tone perturbation. Inflection in Koh, it has been noted, is restricted to tone representing the mood distinction between factive and non-factive. Lexical tone on the verb (low or mid) represents factive mood while high tone represents non-factive. There is a phenomenon in Koh where under certain conditions, factive tone becomes high. Mid tone on an inherently stative verb which is either reduplicated in an affirmative sentence or carries other verbal elements (adverbs, negation) becomes high when it is followed by a pronoun bearing low tone. This is summarized in the following rule:

V -> V /______ PRÒ

Condition: the verb (V) is stative and is either reduplicated, negated, or accompanied by adverbs.

The following paradigm illustrates this perturbation of tone. **koo** is a mid tone verb meaning ´to know´. The perturbation of tone is seen on the reduplicated particle of the verb preceding the first and second person singular and the second person plural pronouns.

(80) a. ka kóo mì koo
he RED me know:F
'He knows me.'

b. ka kóo mù koo
he RED you know:F
'He knows you.'

c. ka koo ni koo
he RED him know:F
'He knows him.'
d. ka koo bburu koo
he RED us know:F
'He knows us.'

e. ka kóo rì koo
he RED you:PL know:F
'He knows you.'

f. ka koo ri koo
he RED them know:F
'He knows them.'

As seen from the paradigm the low tone of the first and second singular and the second plural pronouns has caused the normal mid tone of the verb in the factive to rise to high, although the sense remains factive. The same tone perturbation from mid to high has been noted in two other environments, as now explained.

The verbal affix **-ka** carries a high tone when followed by a low tone pronoun. This is true whether the verb involved is in the main clause (81) or the subordinate clause (82). Examples (81a) and (82a) show the mid tone of **-ka** when it is followed by a pronoun carrying mid tone, while (81b) and (82b) show the high tone **-ká** when it is followed by a low tone pronoun.

(81) a. rúo bburu háŋa ni nàŋ
ą́-ka
tomorrow we give:NONF him millet
'Tomorrow, we will give him some millet.'

b. rúo bburu háŋá rì nàŋ
tomorrow we give:NONF you:PL millet
'Tomorrow, we will give you some millet.'

(82) a. bburu hìh bbày háŋa ni gaŋ ndɨh ròo
we want:F that ive:NONF him wait little first
´We want him to wait a minute.´

b. mì hìh bbày háŋá rì lo sòro
I want:F that give:NONF you:PL talk:F visit

zìŋ bburu
with us

´I want you to visit with us.´

The third environment where tone perturbation has been noted is in two-syllable imperative forms where the verb carries a high tone (i.e., Class 1 low).[13] This high tone is pronounced on the first syllable while the second normally retains its lexical tone. But this lexical tone changes to high before a low tone pronoun. Thus, in (83a) the tone of the second syllable remains unchanged before the mid tone pronoun **ni**, but in (83b) it changes to high before the low tone **mì**.

(83) a. mu úla ni ya
you deceive:NONF him NEG
´Don´t deceive him!´

b. mu úlá mì ya
you deceive:NONF me NEG
´Don´t deceive me!´

In each of the three environments in which tone perturbation has been observed, the tone of the verbal form (whether the entire root, an affix, or the second syllable) has been raised to high when followed by a low tone pronoun.

4.1.5 Summary: verbal inflection. Verbal inflection in Koh is limited to the domain of tone and is used to make one basic mood distinction: factive versus non-factive. Factivity, which refers to the actuality/reality of a given situation, is indicated by base tone on the verb. Base tone, either mid or low, is the lexical tone of the verb and can be most readily seen in a negated factive sentence. The complementary mood distinction of non-factivity refers to the possibility/potentiality of a given situation. It is indicated by a high tone on the verb.

Both factivity and non-factivity are notions that can be further subdivided semantically. In the absence of contextual factors imposed by adverbials or aspectual markers, the inherent aspect of the verb will strongly influence how a factive or a non-factive is understood. Thus, factivity may be variously understood as referring to past completed acts, past continuous, present acts, present continuous, habitual, or general truth, depending on the inherent aspect of the verb which bears the factive marking.

Non-factivity may refer to possibilities, desire or purpose, obligation, contrafactual situations, or hortatives and imperatives, again, depending on the inherent aspect of the verb which bears the non-factive marking. Imperatives are analyzed as non-factive. Although they do not always involve high tone on the verb, there is a raising of the tone on some clausal constituent, and their meaning is non-factive as defined here.

Base tone and high tone are the inflectional forms that appear on the verb. A rising tone sometimes occurs. It has been analyzed not as a third inflectional form but as a combination of factivity and non-factivity on one verb. Under certain conditions tone perturbation from mid to high occurs. In these instances the tone changes without a corresponding change of mood category.

4.2 Auxiliaries and aspect distinctions

In addition to the formal marking of mood through verbal inflection, Koh has access to certain grammatical devices that express formal categories of aspect. Aspectual categories refer to "different ways of viewing the internal constituency of a situation" (Comrie 1976:3). By means of syntactic aspect markers, the speaker may either focus upon the entirety of a situation without referring to its duration or internal phases, or focus upon the internal process of the situation. Formal aspect markers indicate how an event is realized. Aspectual distinctions in Koh are indicated by two auxiliaries (Sect. 4.2) and two aspectual particles (Sect. 4.3). Syntactic aspect interacts with the inherent aspect of the verb. The inherent telicity and/or stativity of a verb affects the interpretation of the aspect markers. In this and the following sections I will examine the formal categories and their meaning as they are combined with inherent aspects of the verb.

4.2.1 ddò and ´PROGRESSIVE´. The auxiliary verb **ddò** indicates progressive aspect. As the main verb of a clause, **ddò** is the locative copula meaning ´to be located´ or ´to stay´.

(84) mì ddò púoo bbi
I LOC house my
´I´m home.´

When **ddò** is used as an auxiliary, its complement is a verb obligatorily bearing a high tone.

(85) mì ddò mbúo hùl
I PROG build:NONF house
´I´m building a house.´

In (85) the low tone of **ddò**, indicating factive mood, asserts that the situation is real and has reached a certain point of development at the time of the utterance. It is also possible for the auxiliary to carry a high tone, as in (86). In this case the entire situation is placed in the realm of the potential.

(86) mì ddó mbúo hùl
I PROG build:NONF house
´I will be building a house.´

The pattern for indicating progressive is as follows:

x + ddò/ddó + V́ + y

The progressive aspect **ddò** marks a subcategory of imperfectivity in that it focuses attention[14] upon the internal temporal constituency of the situation. The use of **ddò** asserts the durativity of a situation, as do all subcategories of imperfective aspect. Progressive aspect is further distinguished within imperfectivity in that it does not imply habituality. It is non-habitual in the sense that the auxiliary **ddò** does not in itself signal an habitual situation, although it can co-occur with an adverbial (phrase) indicating habituality.

(87) mì ddò mbúo hùl bbaŋguoo
I PROG build house always
´I´m always building houses.´

(88) bbày féw lèw tédd, múnu ká póla ná mì ddò
with moon ? cold like as before here I PROG

mbúo hùl
build house

´In January, I´m usually building houses.´

In summary, the primary meanings of the progressive aspect, marked with **ddò**, are durative, non-habitual, dynamic, and progressive.

When **ddò** occurs with verbs which are inherently stative or telic (events), there are resultant changes either in the interpretation of progressive or in the meaning of the verbs. Even with process verbs, which are inherently durative and dynamic, the addition of **ddò** limits the possible rendering of the verb. We examine now the restrictions of effects of combining progressive aspect with the inherent aspectual classes of states, events, and processes.

It would be expected that progressive aspect, which includes a meaning of non-stative, would be incompatible with inherently stative verbs. There are in fact a few verbs which cannot have a progressive form, as in (89).

(89) *mì ddò hím fè
I PROG want thing
*'I'm wanting something.'

Other inherently stative verbs co-occur with the progressive aspect marker, but their meaning is then changed. Thus, in (90a) **koo** may mean either 'to know' or 'to see', but with the progressive in (90b) it can only mean 'to see'.

(90) a. mì koo fè
I know thing
'I see something' or 'I know something.'

b. mì ddò kóo fè
I PROG see thing
'I'm looking at something.'
*'I'm knowing something.'

In (91a) and (91b) the meaning of **tą̀-tąh** changes when it occurs in the progressive construction.

(91) a. mì tą̀-tąh
I RED-tired
'I'm tired.'

b. mì ddò tą́-tąh
I PROG RED-tired
'I'm resting.' ? 'I'm still tired.'

When other stative verbs are put into the progressive construction, this overrides their inherent stativity, and the situation becomes dynamic, as in (92) and (93).

(92) mbìh ddò záŋa
water PROG hot
'The water is becoming hot.'

(93) mì ddò ŋgóoŋ-ŋgooŋ
I PROG strong
'I'm becoming strong.'

With certain stative verbs it is difficult to determine when **ddò** is an auxiliary, marking progressivity, and when it is the locative in a purpose clause in which **bbày** is optionally deleted. When **bbày** is present, as in (94a), **ddò** is the locative copula, but without **bbày** the function of **ddò** is unclear, as in (94b).

(94) a. mì ddò bbày tą́-tąh
I LOC that RED-rest
'I'm staying to rest.'

b. mì ddò tą́-tąh
I LOC/PROG RED-rest
'I'm resting' or 'I'm staying to rest.'

The preferred construction with the lexical class of intransitive stative verbs is **ddò** with **bbày**. This construction is clearly a purpose clause with a locative copula. This indicates that progressivity is marginally compatible with inherently stative verbs.

In summary, then, progressive aspect may not be used with all inherently stative verbs. When it does occur with these, it overrides the verb's inherent stativity.

When progressive aspect is combined with event verbs, the inherent perfectivity of the verb is overridden by the progressive. The situation is thus viewed dynamically and duratively. The verb remains telic, but the 'built-in' termination point has not been reached, as in (95b).

(95) a. ka mboŋ lédd
she birth child
'She gives/gave birth.'

b. ka ddò mbóŋ lédd
she PROG birth child
'She is giving birth.'

Certain event verbs, when combined with the progressive, can be interpreted iteratively (96).

(96) ka ddò íka này
he PROG kill animal
'He is killing animals.'

Even here the emphasis is not upon the individual acts but upon the fact that these acts span a period of time.

Process verbs, which are inherently dynamic and durative, readily accept progressive aspect. Marking a process verb as progressive makes explicit the durativity of the situation and excludes a perfective meaning.

(97) ka ddò rúka sùnum
he PROG plant peanuts
'He is planting peanuts.'
*'He plants/planted peanuts.'

In summary, **ddò** is an auxiliary which indicates progressive, and by implication a situation that is durative, non-habitual, and dynamic. When the progressive appears with an inherently stative (non-dynamic) verb, either the meaning of the verb is changed, or the verb loses its stativity and the situation is viewed dynamically. When combined with an event verb, the progressive aspect causes the verb to lose its inherent perfectivity. Progressive aspect co-bines naturally with process verbs and excludes a perfective interpretation of the verb.

4.2.2 bba and ´HABITUAL´. In discussing the possible interpretations of state, event, and process verbs in the factive mood with no aspect marked, I stated that, as the context allows, the verb may have an habitual sense (Sect. 4.1.1). There is a formal construction in Koh that explicitly indicates habituality. The descriptive copula **bba** ´to be´ (Sect. 3.1) is also used as an aspect auxiliary with the general meaning of ´HABITUAL´. The same construction is found in Karang (Ubels 1983:25-26). But whereas this construction is commonly used in Karang (Ubels, personal communication), in Koh it is limited to contexts relating to giving counsel or instruction to children.

The habitual construction consists of the verb in the factive mood followed by **bba** (glossed HAB) and the same verb in the non-factive mood. If an object is present, it is repeated after **bba** and before the verb in the non-factive (98) and (99).

(98) mù ddah fè bba fè ddáh bugru ya ndáw
you do:F thing HAB thing do:NONF unjust NEG also
´Don´t be doing things unjustly.´

(99) mù bbah bbày bba bbày bbáh ríbba-rì ya
you talk:F word HAB word talk:NONF thief-PL NEG
´Don´t be talking like thieves.´

By using this construction it is possible to form many sentences that are grammatical and refer to an habitual situation. However, they are not commonly used by the Koh. Perhaps the habitual construction had a much wider usage in times past, as it continues to have in Karang, but its usage has narrowed to formal contexts of giving counsel.

4.3 Aspectual particles

In the previous section we have examined the formal aspect categories of progressive and habitual as they are marked by the use of auxiliary constructions. We now turn our attention to a set of two particles which occur at the end of a clause and formally mark an aspectual opposition between imperfective and perfective or non-perfect and perfect. We examine first the particle **bbáy** which is the imperfective marker, although it also carries a non-perfect meaning. Imperfective makes "explicit reference to the internal temporal constituency of a situation" (Comrie 1976:24). The non-perfect indicates that "the situation overlaps an established reference point" (Ubels 1983:29). The second particle of this set is **ró**; it combines perfective aspect with the perfect. Perfective aspect views the situation in its entirety without reference to the internal temporal constituency. The perfect "tells us nothing directly about the situation itself, but rather relates some state to a preceding situation" (Comrie 1976:52).

4.3.1 bbáy and ´IMPERFECTIVE´. Part of the meaning of **bbáy**, ´imperfective aspect´, is that the situation is "unfolded into its component parts" (Comrie 1976:23). When **bbáy** is used the situation is being viewed imperfectively. **bbáy** is usually translated as ´still´ or ´yet´. In (100) we see **bbáy** with a process verb and in (101) with a stative.

(100) ka ru sùnum bbáy
he plant:F peanuts IMPFT
´He is still planting peanuts.´
*´He planted peanuts.´

(101) mì hɨ̀-hɨee bbáy
I RED-fear:F IMPFT
´I´m still afraid.´ *´I was afraid.´

The situation is durative, but differs from progressive in that the situation is not necessarily dynamic. It may be a continued state as seen in (101). **bbáy** is not simply continuous since the situation may be repetitive. Example (100) could also be translated as, ´He is planting peanuts again.´ When an event verb, which is inherently perfective, appears with the imperfective particle **bbáy**, the resultant meaning is often repetitive, as in (102) and (103).

(102) ka mboŋ lédd bbáy
she birth:F child IMPFT
´She is having a child again.´

(103) ka mbuo hùl bbáy
he build:F house IMPFT
´He is building again.´

bbáy, therefore, represents imperfectivity. It refers to a state or action which is continuous, durative, and in some instances repetitive. It can co-occur with either the factive or non-factive mood. Lacking any further specification from the context, **bbáy** appearing with the non-factive mood is usually interpreted as referring to a continued state/action in the future (104) and (105).

(104) ka mbúo hùl bbáy
he build:NONF house IMPFT
´He will be building (other) houses again.´

As seen in (105), this future continued state/action may either be the continuation of a present state/action into the future or be another instance of the state/action started in the future.

(105) ka ddí sìm bbáy
he sing:NONF song IMPFT
´He´ll be singing again´ or ´He´ll still
be singing.´

Imperfectivity, then, is the principal meaning of the aspect particle **bbáy**. **bbáy** also has a non-perfect aspectual component to its meaning, in that it locates the situation in temporal reference to some external time point.[15] Thus, **bbáy** functions in opposition to the particle **ró** (cf. Sect. 4.3.2). In marking the non-perfect, **bbáy** is used in relation to an established reference point that may be the moment of speech, a time in the past or future, or even a completely relative time, as seen in the following examples:

(106) Moment of speech:

púoo ddò bbáy kère bbáy
village LOC with good IMPFT
´The village is still fine.´

(107) Relative time:

bbe rŏo, mù kăwke waqa bbáy
that then you rest a:little IMPFT
´With that (by doing that) you will be resting secure.´

(108) Future:

nàŋ, à gáŋ bbà póla bbáy
millet it grow:NONF to advance IMPFT
´The millet will continue to grow.´

The particle **bbáy**, then, combines the aspectual categories of imperfective and non-perfect.

4.3.2 ró and ´PERFECT´. Where **bbáy** marks imperfective and non-perfect, **ró** marks perfective and perfect. In marking perfectivity, **ró** is completive and presents the situation as a whole, without reference to any internal phases of that situation. It refers to a "complete situation with beginning, middle, and end" (Comrie 1976:18). In the following sentences the entire act is in focus and not simply the termination of the act.

(109) mì mbuo hùl ró
I build house PERF
´I´ve already built a/the house.´

(110) ka mboŋ lédd ró
she birth child PERF
´She has already had a/the child.´

The perfective form may in fact indicate the completion of an act/state, as when an inherently perfective form in the factive mood with the particle **ró** indicates a fully developed situation in the past. But there are instances when the perfective form actually has an inceptive meaning.

This is the case with certain stative verbs in the factive mood. Example (111) illustrates this.

(111) mì hɨee ndày ró
I fear:F cow PERF
´I began to fear cattle.´

While the perfective of some stative verbs indicates entry into that state, for other statives it describes a state which has been fully developed in the past, as with the verb **kèr** ´to think´ in (112).

(112) mì kèr bbày ró
I think:F word PERF
´I´ve already had an idea´ or ´I´ve already thought of that.´

When the perfective marker **ró** is combined with a process verb in the factive mood, the resultant meaning is that an instance of that verb has been accomplished in the past. In (113) the entire act is past.

(113) ka ddaŋ mbìh ró
she draw:F water PERF
´She has already drawn water.´

It is more difficult to ascertain the sense of **ró** when it is combined with the non-factive mood. In a simple clause I have not found an instance where non-factivity and **ró** co-occur on stative verbs. With event and process verbs, high tone and **ró** seldom appear together in a simple clause. When they do it is with the sense that an instance of the given situation will certainly recur. As noted below (114, 115), **ró** does co-occur with verbs in the non-factive mood in certain types of complex clause constructions.

In each of the examples given above, **ró** not only refers to the whole of a situation, but it also refers to that situation as occurring prior to a specific reference point of time but still affecting that reference point. The reference point above is the moment of speech. Thus, in (109) the house still stands, in (110) she still has a child, in (111) I still fear cows, etc. **ró** in this manner marks the perfect. When the point of reference is the moment of speech, **ró** expresses the relation between the present state and the past situation. As the perfect marker, **ró** can also mark the relation between a past state and an even earlier situation or between a future state and an earlier situation (Comrie 1976:53). It is in the last relationship that we often find **ró** appearing with non-factivity. In (114) and (115) the action of the main clause modified by **ró** will occur in the future, but before the action expressed in the subordinate clause.

(114) zàddká mì yɨ̀h, mí vi lee, Serafine,
when I return:F I come:F CONJ Serafine

à zíŋ lédd-pyaŋ ró
she find:NONF child-baby PERF

‘When I return, Serafine will already have had her baby.’

(115) zàddká mù ddò kér bbày lee, bbi, mì bbáh
when you PROG think word CONJ me I talk:NONF

bbày ró
word PERF

‘When you are still thinking, I will already have finished talking.’

In summary, the particle **ró** combines the aspectual notions of perfective and perfect.

5 Complex constructions: complement-taking verbs and serial verbal constructions

Having described the formal TAM categories of Koh in Sect. 4, I turn now to the consideration of two complex verbal constructions that occur very frequently: complement-taking verbs and serial verbal constructions. In Sect. 5.1 I examine the two types of complement-taking verbs and the sentence constructions in which they occur. In Sect. 5.2 serial verbal constructions are examined.

5.1 Complement-taking verbs

There are two types of complement-taking verbs: those taking object complements and those taking complements that are not objects. Object complement-taking verbs have an overt subject which is the agent of the complement-taking verb. Non-object complement-taking verbs have no nominal functioning as subject. The verbal elements (aspectual particles, adverbs, and the negation marker) are obligatorily clause final in object complement-taking verbal constructions, but may occur in non-final position in non-object complement-taking constructions. These distinguishing features can be seen in the examples that follow in this section.

5.1.1 Equi-construction. In this first type of construction, which is restricted to object complement-taking verbs, there is no overt subject in the complement clause, but the subject of that clause is understood to be co-referential with the subject of the matrix clause. The subordinate verb is always in the non-factive mood, indicating a potential but unrealized situation in relation to the complement-taking verb in the matrix clause. In (116) and the following examples the complement clause is introduced with the complementizer **bbày** ´that´.

(116) mì hìh bbày séke zìŋ ni
I want:F that go:NONF with him
´I want to go with him.´

Aspectual particles, adverbs, and the negation particle are clause final. It may occasionally be unclear which verb is being modified, as illustrated in the glosses of (117).

(117) mì hìh bbày séke zìŋ ni ròobbáy
I want:F that go:NONF with him again
´I want again to go with him (may never have gone before)´ or ´I want to go with him again.´

5.1.2 Manipulative with bbày. When the speaker/subject desires another person to do something, another construction is used as follows:

(NP) + V + bbày + há + NP + V +/- NP +/- x
COMP give:NONF +F

Either an object complement-taking verb (118) or a non-object complement-taking verb (119) may occur in the matrix clause.

(118) mì hìh bbày hạ́ mù kaw
I want:F that give:NONF you stay:F
´I want you to stay.´

(119) ndàdd bbày hạ́ mù kaw
good:F that give:NONF you stay:F
´It´s good that you stay.´

In a sentence using a complement-taking verb the speaker/subject is attempting to influence or manipulate a situation over which he may have little or no control. This may account for the indirectness manifested in the subordinate clause, which means literally, ´give/cause or let x do y.´ **hạ** in an independent clause is the verb ´to give´ and is used here as a form of causality. We will see in Sect. 5.2.1 that **hạ** is often used in Koh in conjunction with another verb with the meaning of ´permit´ or ´oblige/make´. When **hạ** functions as the main verb of the subordinate clause, it is always non-factive while the second verb is factive. Although it is rarely seen, it is possible to omit **hạ** and use only the other verb, which is then always non-factive. Example (120a) is the more frequent construction with **hạ**. In (120b) **hạ** is omitted.

(120) a. ndàdd bbamba bbày hą́ mù bbo
good very that give:NONF you counsel:F

vu-oo
son-your

´It´s good that you counsel your son.´

b. ndàdd bbamba bbày bbőko vu-oo
good very that counsel:NONF son-your
´It´s good that you counsel your son.´

In the subordinate clause the object of **hą** is always the same referent as the subject of the second verb. If the object of **hą** in the subordinate clause is the second person singular (**ni**) or plural (**rì**) or the third person plural (**ri**) pronoun, a problem arises since these pronouns have separate forms for the subject or object. We find in this case that the object pronoun is used and that the corresponding subject pronoun may optionally follow, as in (121) and (122).

(121) à ndắdd bbamba bbày hą́ ni
and good:NONF very that give:NONF her

(à) vi
she return

´It will be important that she return.´

(122) liee, mì vi mí kèr bbày háŋa ri (i)
yesterday I come I desire that give them they

kpà (i) kpà mì làri
loan they loan me money

´Yesterday I came, wanting them to loan me money.´

Since the subordinate clause verb is always non-factive, the temporal reference point of the verb is ambiguous unless specified by a temporal adverb or phrase. When the action expressed in the subordinate clause is subsequent to the time of the complement-taking verb, temporal phrases can be used to specify the relationship. In (123), for example, the desire expressed by the complement-taking verb is present, but the desired action is for a future time. The action expressed by the complement-taking verb may also be unrealized at the moment of speech and the action desired be at an even further unrealized time, as in (124).

(123) mì hìh bbày há mù vi rúo
I want:F that give you come tomorrow
´I want you to come tomorrow.´

(123) mì hìh bbày há mù vi rúo
I want:F that give you come tomorrow
´I want you to come tomorrow.´

(124) ka vbíka mù tílaw kéy bbày há mù se
he ask:NONF you evening this that give you go

nzáh-mbìh rúo
river tomorrow

´He will ask you this evening to go to the river tomorrow.´

Whereas verbal elements only occur clause finally with object complement-taking verbs, they may occur immediately after non-object complement-taking verbs (125a) or sentence finally (125b), depending on the intended meaning.

(125) a. ndàdd bbamba ròbáy bbày hą́ rì zùo zìŋ kɨ
good very still that give them place with REFL

zìŋ kɨ
with REFL
´It´s still good that they agree together.´

b. ndàdd bbamba bbày hą́ rì zùo zìŋ kɨ
good very that give them place with REFL

ròobbáy
still

´It´s good that they still agree together.´

5.1.3 Manipulative with zàddká. This construction occurs only with non-object complement-taking verbs and is formed as follows:

V + zàddká + Complement Clause

zàddká is a subordinating conjunction which may be translated as ´if´, ´when´, or ´that´. It can introduce a conditional clause (126), a temporal situation (126), or a general truth statement (127).

(126) Conditional or temporal:

ndàdd bbamba zàddká mù se hùl są̀w-pùh
good very CONJ you go:F house root-wood
´It is good when/if you go to the dispensary (specific situation).´

(127) General truth:

ndàdd	bbamba	zàddká	máy-rì	ddah	mbìh	sì	hùh
good	very	CONJ	woman-PL	do:F	water	on	fire

'It's good that women boil water.'

In this construction there tends to be agreement of factivity/non-factivity between the complement-taking verb and the subordinate verb. Thus, in (128) both verbs are non-factive, whereas in (126) they are both factive.

(128)

à	ndádd	bbamba	zàddká	mù	séke	hùl	sạ̀w-pùh
and	good:NONF	very	that	you	go:NONF	house	root-wood

'It will be good/better when you go to the dispensary.'

There are exceptions to this tendency, as in (129).[16]

(129)

à	ndádd	bbamba	zàddká	ka	yɨ̀h,	à	vi
and	good:NONF	very	CONJ	she	return:F	she	come:F

'It will be important that she return.'

5.1.4 Explicit benefactor and ká. In contexts other than with complement-taking verbs **ká** is used straightforwardly as a relativizer, introducing either a subject relative clause (130) or an object relative clause (131).

(130)

ka	bba	lédd	ká	kàw	zìŋ	yụ-míee
she	DCOP	child	REL	stay	with	brother-mother:her

'She was a child who stayed with her uncle.'

(131)

mbána	ká	mì	mgba	ku	ná,	mì	ddah	pédd...
engagement	REL	I	take	there	here	I	do	work

'The engagement that I entered, I worked. . .'

There is a special usage of **ká** with complement-taking verbs, where **ká** is used rather as a complementizer 'that' to introduce a subordinate clause. The matrix clause includes either a subject with an object complement-taking verb or a non-object complement-taking verb and a benefactive phrase. The function of this construction is to make explicit the beneficiary of the desired action of the subordinate verb.

(132) ndàdd bbamba bbày púo ká mì se ddi
good very for village COMP I go:F there
'It's good for the village that I go there.'

It is possible for the beneficiary to be the subject of the subordinate clause, in other words, the agent who performs an action from which he is the beneficiary (133).

(133) ndàdd bbamba ró bbày nzòbb-rèw-púo-rì
good very PERF for person-eat-village-PL

ká i mạy bbày zìŋ kɨ
COMP they talk word with REFL

'It's good for the authorities that they discussed together.'

Contrary to what is seen in the other complement-taking verbal constructions, in the case of these **ká** complements the verb may be factive or non-factive, and such mood marking is based solely on the semantic-pragmatic factors relating to the **ká** clause itself, independent of the factivity marking in the matrix clause. Thus, as (132) above indicates a realized situation, (134) indicates a potential situation.

(134) ka hìh bbày púo ká mì séke bbà bbàddi
he want for village COMP I go:NONF to there
He desires for the village that I would go there.'

5.2 Serial verbal constructions

Serial verbal constructions (henceforth SVC) are another phenomenon occurring in the Koh verbal system. A. Bambose (1974:17) defines SVCs or "serial verbs" as a term applied "to the combination of verbs found in many West African languages where all the verbs share a common subject in the surface structure." An SVC is a construction in which two or more verbs appear within a single simple clause in which there is only one subject. In Koh, the first verb in an SVC carries the TAM distinctions while the following verb(s) is/are marked as factive by the base tone.[17] While in some African languages there is a wide range of verbs which appear in SVCs, the number of verbs in Koh which admit serialization with other verbs is restricted. It is thus feasible to state a closed list of verbs which occur in serial constructions.

On the basis of semantic distinctions, it is possible to speak of two basic types of SVCs in Koh. These two types are referred to by Bamgbose (1974:18) as **linking SVCs** and **modifying SVCs.**[18] In linking SVCs each verb retains its own meaning (although this meaning may be modified slightly), and there is present in the verbal string the idea of simultaneity, sequence, and/or consequence. Modifying SVCs, on the other hand, lack sequentiality or consequentiality.

Modifying serial verbs function "merely to modify the meaning of the other verbs with which they occur" (Bamgbose 1974:33). In the following two sections (5.2.1 and 5.2.2) we will examine the principal verbs which occur with these two types of SVCs and which exemplify the types of constructions which are found within each type of SVC.

5.2.1 Linking SVCs. In this section I examine three verbs that occur frequently in linking SVCs. These are **hạ** ´to give´, **vi** ´to come´, and **vǎ** ´to go´.

5.2.1.1 hạ ´to give´ - ´benefactive/causal´. In a simple clause **hạ** as a main verb means ´to give´, as seen in (135).

(135) mì hạ mù fè
I give you thing
´I give/gave you something.´

In SVCs **hạ** may be used to indicate a benefactive relationship. In this function **hạ** is always the second verb in the verbal string, as seen in (136) and (137).

(136) ka zùo ngáw hạ bburu
she make food give us
´She is making food for us.´

(137) i yạh máy hạ ni
they take wife give him
´They took (found) a wife for him.´

This same construction is found in sentences with bitransitive verbs when the indirect object becomes the direct object. Example (138a) shows the normal order in bitransitive clauses. Example (138b) shows the corresponding clause using **hạ**.

(138) a. mù kiee mì nzáh Koh
you teach me language Koh
´You are teaching me Koh.´

b. mù kiee ni hạ mì
you teach it give me
´You are teaching it to me.´

We have already noted in the discussion on complement-taking verbs (Sect. 5.1.2) that **hạ** can be used to indicate causality. **hạ** also has this function in serial verbal constructions of consequence. When one action results in another action, this is often expressed in Koh with an SVC using **hạ** and another verb.

(139) lɛɛ, làw-ee vǐ́ fạh ni pǘh hạ ni zo
and heart-his come anger him a:lot give him jump

zo kuruk
jump far

´And becoming very angry, he jumped a long way.´

(140) mù se, mú vǎ yạh fè nzòbb hạ mboolke
you go you go take thing person give debt

ka ddò túl-a ya
it LOC head-your NEG

´Don´t go take something from someone so that a debt is over your head.´

hạ appearing in an SVC, then, has the function either of indicating a benefactive relationship or of indicating causality.

5.2.1.2 vi ´to come´. A very common SVC involves the verb **vi** ´to come´ followed by another verb in a sequential relationship. Sequentiality of action is the interpretation of **vi** when it is followed by an active verb, either an event (141) or a process (142).

(141) mì vi mbuo hùl
I come build house
´I came and built a house.´

(142) mì vi ru sùnum
I come plant peanuts
´I came and planted peanuts.´

When **vi** is followed by a stative verb, two interpretations are possible. First, if the idea of movement is still present, the second verb can describe the state of the subject who is coming. Thus, in (143) the state of being tired is simultaneous to the action of coming.

(143) mì vi tạ̀-tạh
I come tired
´I came tired´ or ´I was tired when I came.´

Second, with verbs describing mental activities/perceptions, when the idea of movement is excluded, **vi** is used to describe a recent past occurrence. The semantic content of **vi** meaning a movement in space toward a location is metaphorically extended to mean a movement in time toward a moment.

(144) ka vi yìŋnra
he come forget
´He just forgot.´

(145) pạra nún-i ví koo fè
seed eye-my come see thing
´My eyes are just starting to see (understand).´

In SVCs with **vi** the point of time of reference is the location of the speaker or the moment of speech, unless otherwise specified.

5.2.1.3 vǎ ´to go´. The verb **vǎ** contrasts with **vi** in SVCs. While **vi + V** means ´come and do´, **vǎ + V** means ´go and do´. SVCs with **vǎ** are illustrated with an event (146) and with a stative verb (147).

(146) ka vǎ mbuo hùl
he go build house
´He left and built a house.´

(147) mì vǎ hɨ̀-ɲɨee
I go fear
´I left and was afraid.´

The verb **vǎ** here represents a displacement in space from the location in which the statement is made to or towards another place. It is difficult to find a statement in which the idea of movement is not present, as evidenced by the fact that **vǎ** is interpreted spatially even with most stative verbs, (147). When the idea of movement is not made explicit or is excluded, **vǎ** takes on a secondary meaning of immediacy in the future. Here it represents a displacement in time from the moment of speech to or towards another moment. In (148) **vǎ** may mean ´go´ or ´immediacy´.

(148) ka vǎ ru sùnum
he go plant peanuts
´He is leaving to plant peanuts´ or ´He will
plant peanuts.´

The status of **vǎ** as a verb is problematic in that it never occurs as an independent verb in a simple clause. The active verb form **se** ´to go´ is used in simple clauses to depict movement away from a source. In spite of this, **vǎ** is interpreted as an anomalous verbal form in SVCs for the following reasons:

The form **se** never occurs in an SVC. It is always replaced by **vǎ**. **se** and **vǎ** may co-occur in a multiple clause (cf. Sect. 1.5.2.2).

(149) mì se, mí vă ddaŋ mbìh
I go I go draw water
'I went and drew water.'

vă is rejected as an auxiliary form since the following verb carries a factive (base) tone, whereas with known auxiliaries the following verb is always non-factive.[19] For the same reason **vă** cannot be part of the modal system (marking likelihood/near future possibility), since one would in such a case expect **vă** to be combined with high tone/non-factivity. In addition, if **vă** were to be a particle representing likelihood in the modal system, one would expect it to be able to modify any verb of movement. In fact, it cannot co-occur with verbs of movement toward something. This same fact is easily explained if **vă** is a verb indicating movement away from something.

vă can also not be interpreted as part of the aspectual system, since it can occur with any of the established aspect markers.

(150) a. ka vă mbonŋ lédd ró
she go birth child PERF
'She already left to have a child.'

b. ka vă mboŋ lédd bbáy
she go birth child IMPFT
'She left again to have a child.'

c. ka vă ddò kér bbày
he go PROG think word
'He left thinking.'

Thus, **vă** cannot be accepted as an auxiliary or part of the modal or aspectual systems. In most instances **vă** carries the meaning of directional movement away from something. In cases where movement is excluded as a possible sense of **vă**, it carries the idea of immediacy in the future, parallel to the usage of **vi** as recent past. **vă** functions as a replacement for **se** 'to go' in serial constructions. For these reasons **vă** is interpreted as a verb that occurs in linking SVCs.

5.2.2 Modifying SVCs. Unlike linking constructions in which each verb retains its meaning in the "linking" or combining of two or more actions, modifying verbs in SVCs cannot be interpreted according to their meaning as main verbs in a simple clause. In the modifying SVC the sentence has only one verbal action which is modified by a second verb. In this section two verbs that commonly occur in modifying SVCs are considered: **òo** 'to be finished' (5.2.2.1) and **màh** 'to be sufficient' (5.2.2.2). They are representative of a small class of verbs that occur in typically serial constructions with the function of modifying another verb. I have analyzed these forms as verbs, although they may be in the process of losing their verbal characteristics and beginning to function more like adverbs. A third construction involving the verb **ddah** ´to do´, which has some of the characteristics of a modifying SVC, is considered in Sect. 5.2.2.3.

5.2.2.1 òo ´to be finished´. **òo** in a simple clause is the verb ´to be finished´. Used in an SVC as a post-verbal modifier, it indicates the termination of the action of the verb.

(151) i sa gbána òo ríw bele
they extend net finish all together
´They all finished extending their nets.´

(152) ka kàh hùl bbe ná òo ya bbáy
he roof house his there finish NEG IMPFT
´He hasn´t yet finished roofing his house.´

When **òo** is used in an SVC, there are no two other phrases that would have the equivalent meaning of the one sentence with the SVC. That **òo** simply modifies the action of the first verb and cannot be regarded as a separate clause is confirmed in that negation of the modifier (as in 152) is in fact negation of the SVC. It would be impossible to negate the verb and retain the sense of **òo** ´finish´.

5.2.2.2 màh ´to be sufficient´. Another post-verbal modifier is the verb **màh**, which as a main verb means ´to be sufficient/enough´ or ´to be able´, as seen in (153).

(153) pédd màh mì ró
work:N sufficient me PERF
´I have sufficient work already.´

When **màh** appears in an SVC following the main verb, it must have an overt object, although **màh** itself modifies the preceding verb. As with **òo** above, negation of the modifier **màh** is sentence negation.

(154) mbàm sɨ́wni too màh fè ya hóy
rain:N year:this rain:V enough thing NEG only
´It hasn´t yet rained enough this year.´

The SVC in which **màh** occurs can be used in Koh to express equative, and with the negative particle to express the comparative ´less than´. In both cases **màh** is followed by a nominal.

(155) ka su̩ fè màh mù
he eat thing equal you
´He eats as much as you.´

(156) ka ddah pédd màh biee ya
he do work equal father:his NEG
'He works less than his father.'

The verb **màh** co-occurs with **fè** 'thing' in an SVC to express the sufficiency of an action, as seen in (154) and the following examples:

(157) ka kèr bbày màh fè
he think word enough thing
'He thinks enough.'

(158) ka ddò bbáy wáh màh fè
he LOC with field enough thing
'He has enough fields.'

(159) zàddká à ŋgooŋ màh fè lee, à sé kɨni
when he grow enough thing CONJ he go hunting
'When he is big enough, he'll go hunting.'

In each of these SVCs **màh** does not have an active verbal sense but is merely serving to modify the action described by the first verb.

5.2.2.3 ddah 'to do'. In a simple clause **ddah** as a main verb means 'to do', as seen in (160).

(160) mì ddò ddáh pédd
I PROG do work
'I am working.'

When **ddah** appears with another verb in a single clause, the clause has the following construction:

NP + ddah + NP + V

In this construction **ddah** carries the TAM distinctions, and the second verb is always factive (161a) and (161b).

(161) a. mì ddáh gári ŋgadd
I do:NONF clothes hang:up:F
'I will hang up the clothes.'

b. *mì ddah gári ŋgádd
I do:F clothes hang:up:NONF

There are certain verbs in Koh which obligatorily occur with **ddah** when the action is not reflexive with the subject. Examples are **ddah. . .ndùo** 'to bury', **ddah. . .mu** 'to hide', **ddah. . .zuoo** 'to place', and **ddah. . .ŋgadd** 'to hang up'.

(162) a. bburu mu-mu
we RED-hide
'We hide ourselves.'

b. bburu ddah ndàw-hùh mu
we do trap-fire hide
'We hide the weapons.'

(163) a. bburu ndùo lúoo
we enter hole
'We enter the hole.'

b. bburu ddah pùh ndùo lúoo
we do stick enter hole
'We bury the stick.'

Syntactically, **ddah** does not appear to function as an auxiliary in this construction, since the following verb is marked for the factive mood. Likewise, it does not appear to be a complement-taking verb, since it is never followed by the complementizer **bbày.** In addition, the complements of **ddah** in this construction are very different structurally from complements of complement-taking verbs.

As with the modifying serial verbal constructions discussed above, **ddah** does not retain its own verbal meaning. But it does not have a modifying function that is similar to **òo** or **màh.** It functions, rather, as a causative with inanimate objects and with only a limited set of verbs. The linking SVC with **hạ** (Sect. 5.2.1.1) is the primary causal construction in Koh.

ddah also differs from **òo** and **màh** in that it is the first verb in a clause, while **òo** and **màh** both occur after the verb they modify. Bamgbose (1974:34,36) makes the observation, though, that modifying verbs may either precede or follow another verb in serial constructions. He cites several African languages that have both types of modifying verbs. I do not have sufficient data in Koh to establish whether or not other "pre-verbal" modifying verbs occur.

It will be necessary to investigate further the use of **ddah** with a second verb before making a final conclusion as to its status. Because it has certain syntactic and semantic properties that are characteristic of modifying serial verbal constructions, I have made the preliminary analysis that **ddah** here is acting as a verb in a modifying SVC.

5.2.3 Summary. While serial verbal constructions occur in Koh, they are less numerous and less complex than in some African languages. Several of the main SVCs which occur frequently have been discussed above. Both linking and modifying SVCs occur in Koh. Linking SVCs which occur frequently have been discussed above. Both linking and modifying SVCs occur in Koh. Linking SVCs of high-frequency occurrences are those which include verbs carrying meanings of beneficiary, causality, and direction (both "coming" and "going"). Post-verbal (**òo** and **màh**) modifying serial verbs occur. Pre-verbal **ddah** also occurs in a serial construction, where it functions as a causative. In modifying SVCs there is only one action expressed in the clause, whereas in linking SVCs two or more actions are linked together.

6 Summary of verbal forms and their primary meanings

In the analysis of the Koh verbal system, I have attempted to describe the grammatical forms used and their syntactic properties, and to define the basic functions or meanings of each form. I have also attempted to describe some of the interdependencies among inherent aspectual properties of the verb, mood categories, and aspect categories. The base of study has been the sentence. I have emphasized TAM categories in the verbal system and two complex verbal constructions: complement-taking verbs and serial verbal constructions.

This study should provide a foundation for further study. There are two areas of the verbal system that require further investigation: a) the interrelationships of TAM categories between themselves and with other verbal and non-verbal categories, and b) the systematic study of verbal forms and constructions in discourse.

In speaking of the categories of tense, aspect, and mood, Wallace (1982:202) states, "It is an arbitrary division of verbal semantics into compartments which are not quite as easily separable as one is led to believe. Time, aspectuality, and modality--the semantic fields to which the formal categories of tense, aspect, and mood are supposed to refer--are almost inextricably scrambled together." In categorizing forms as "aspect" or "mood," I am stating that their most salient features are within these categories. There is no question that there is considerable overlap among time, modality, and aspectuality in Koh. Much more study needs to be done to identify the semantic properties of verbal categories and their interaction with each other.

In this paper I have described the semantic and syntactic functions of verbal categories independently of their pragmatic functions in discourse, since I haven´t yet systematically studied discourse structures. I agree with Wallace (1982:207): "Verbal categories are important components in the structure of discourses, and indeed, that one does not truly understand ´the meaning´ of a verbal category in a particular language unless one understands its place in discourse." The analysis of the verbal system which has been presented here is a base from which a discourse perspective may proceed.

In what follows, I summarize the verbal categories and their basic meanings and interrelationships in order to provide an overview of the Koh verbal system.

6.1 Terms and definitions of Koh verbal categories

6.1.1 Inherent aspectual classes

States: non-dynamic, durative situations
Events: viewing a situation perfectively; limited in this study to telic situations, whether these are punctual or durative; dynamic
Processes: viewing a situation imperfectively; dynamic; durative; atelic

6.1.2 Mood categories

Factive: a real situation--fully developed or at a stable point of development
Non-factive: a possible or potential situation

6.1.3 Aspect categories

Progressive: marking a situation as dynamic, durative, and non-habitual
Habitual: marking a situation which is characteristic of an extended period of time
Imperfective: marking a situation as continuous, durative, and in certain cases repetitive; includes as a special case habituality; includes the temporal notion of non-perfect
Perfect: viewing the situation in its entirety, completive; includes the temporal notion of perfect

6.2 Interaction of mood and aspect with states, events, and processes[20]

The following three charts show some of the interpretations given to states, events, and processes when they are combined with formal aspect and mood categories.

STATES

	FACTIVE (BASE TONE)	NON-FACTIVE (HIGH TONE)
UN-MARKED	Primary meaning: a present reality; a stable state Other: past state, a general truth	Primary meaning: possibility (often with a future meaning) Other: desire/purpose (following complement-taking verbs); obligation; contrafactual in the past; hortative; imperative
PROG **ddò**	Incompatible with certain states (e.g., **hìh** 'to want') Changes the meaning of certain verbs to denote a process (+dynamic) Cancels the stativity of certain verbs	
IMPFT **bbáy**	A continued state Other: coincides to a point in time	Unrealized A continuous state in the future a) Continuous from the present b) Continuous from an unrealized time
PERF **ró**	Inceptive (i.e., entry into a state still in effect at the moment of speech) Other: fully developed state in the past (-durative) Precedes a point in time	No forms found

EVENTS

	FACTIVE (BASE TONE)	NON-FACTIVE (HIGH TONE)
UN-MARKED	Primary meaning: an accomplished fact (i.e., past action) Other: present action; habitual; general truth (-perfectivity) Special: imminent event	Primary meaning: possibility (often future) Other: (cf. stative non-factive)
PROG **ddò**	Causes the verb to lose its inherent perfectivity; the situation becomes durative, and the termination of the event hasn´t occurred (-perfectivity)	
	Iterative for certain verbs Realized	Unrealized
	Repetitive	Future repetition of an act that has occurred before
	Other: an action in process	Other: certainty that the action will recur
IMPFT **bbáy**	(-perfectivity) Coincides to a point in time Special: with **ddò**, ´is still V + ing´	
	Completed act in the past Other: indicated completion of termination point	"Future perfect" in a co-ordinating clause (i.e., ´will already have + V´)

PERF **ró**	precedes a point in time Special: with **ddò**--inceptive and progressive, ´has started V + ing´

PROCESSES

	FACTIVE (BASE TONE)	**NON-FACTIVE (HIGH TONE)**
UN-MARKED	Primary meaning: either a past or present continuous situation Other: past action; present action; general truth; habitual	Primary meaning: possibility (often future) Other: (cf. stative non-factive)
	Excludes a perfective meaning	
PROG **ddò**	Realized Continuous action (current or in the past)	Unrealized Future continuous action a) Continuous from the present b) Continuous from a future point
IMPFT **bbáy**	Other: repetitive	
	Coincides to a point in time	
PERF **ró**	An instance of the verb has occurred in the past (-imperfectivity) Other: precedes a point in time	"Future perfect" in a co-ordinate clause (-imperfectivity)

Notes

1. The phonological analysis of Koh is being done by my colleague, Ms. J. Camburn.

2. By saying that Koh is an SVO language, I mean that it has an SVO basic clause order and that it has most of the word order properties of VO languages as established by Venneman and discussed by Comrie (1981:80-96): that is, Koh has prepositions, genitive following the noun, modifiers following the noun, and pre-verbal auxiliaries.

3. The simplest way of finding the most unmarked form of the verb in Koh is in the negated factive sentence because a verb root cannot be reduplicated in a negated clause. In order to establish a verb's lexical tone in an affirmative sentence, it is necessary to know where the base tone occurs in a reduplicated verb.

4. Ubel (1983).

5. **-ka** is subject to vowel harmonization by the following rule:

```
-ka  -> kV        /  (C) V       (C)
        -hi             -hi
         bac             bac
         lo              lo
```

In addition, since nasalization is a word level function, **-ka** becomes **-ŋa** when it is affixed to a nasalized verb root.

6. A small class of bitransitive verbs (+0,+IO) are included within the discussion of transitive verbs since they function in an identical manner within the verbal system.

7. It is clear that reduplication has more than one function, since it is also possible to retain an object and reduplicate the verb. Thus, **mì sụ fè sụ** 'I'm eating something.' However, that it has a grammatical function is indicated in that an object cannot be deleted without reduplication taking place.

8. Strictly speaking, inherent aspectual distinctions refer not simply to a verb, but to a verb and its complements, i.e., to a situation. For example, a verb may in certain situations be telic ('to plant a tree') and in others be atelic ('to plant peanuts'). When I refer in this paper to the inherent aspect of the verb, it should be understood that the verbal complements are specifically chosen so as to insure the given aspect.

9. For the formation of these three classes in Koh I have

benefitted greatly from the work of Ubels (1983). We have both been guided by Comrie (1976).

10. The verb **koo** can also mean ´to see´. I´ve chosen to use **koo** in examining the relationships between inherent and syntactic aspect since certain syntactic markers require **koo** to be interpreted non-statively.

11. As Welmers (1973:345) points out, ´He plants´ is not an acceptable English phrase in and of itself. One must say, ´He is planting´ or ´He plants every year at this time.´ The difficulty is in finding an English phrase for an acceptable concept in Koh that distinguishes continuous from present where English does not, at least for this type of verb. Koh does have a distinct aspect marker which excludes a simple present action rendering. But a verb in the factive mood without explicit aspect markers can have either a continuous or a "point of action" meaning.

12. In contrast to the Koh situation, in Karang all verbs, whether the base tone is low or mid, form the imperative by raising the tone of the pronoun to high without modifying the base tone on the verb. In addition, non-factive formation in subordinate clauses with a pronominal subject follows the same pattern (Ubels, personal communication). In both Koh and Karang the pronoun is optionally deleted. In this case the tone of the verb is still guided by the rules of imperative formation.

13. There is only one phonemic tone per morpheme, on the first syllable, and this is marked in the orthography. The change in tone caused by perturbation is marked in the text of this paper, but it does not appear in the orthography, since it is predictable.

14. In the following discussion of aspectual categories in Koh, we proceed from a basic classification of aspectual oppositions as given by Comrie (1976:25):

Perfective Imperfective

Habitual Continuous

Non-progressive Progressive

15. We are accepting here the position and reasoning of Comrie (1976:52f.) that the perfect/non-perfect category may be called an aspectual category "in a rather different sense" from other aspects already discussed.

16. Example (129) is also the construction used for the formation of contrary-to-fact conditions in the past. Thus, it could have the meaning, ´It would have been good if she

has come.´ I lack specific examples in my data to confirm this analysis.

17. In a purpose clause of the structure, **bbày + V́**, the **bbày** may be optionally deleted giving a construction in which two verbs co-occur in one sentence with the second verb carrying a high tone. Purpose clauses with **bbày** deleted may be distinguished from SVCs grammatically by the presence of this high tone and semantically by the meaning of purpose which is always present in purpose clauses but absent in SVCs.

18. Bamgbose distinguishes these two types grammatically by saying that linking SVCs are derived from two or more underlying sentences while modifying SVCs are derived from a single underlying sentence.

19. **vǎ** does occasionally occur in purpose clauses with **bbày** deleted, as described in note 17. In this case, the verb following **vǎ** is non-factive. In all other occurrences of **vǎ + V** the second verb is factive.

20. In the summary charts the habitual aspect is not considered because of its restricted use in the language and the lack of examples of habitual aspect with various types of verbs.

Appendix A Conventions and Abbreviations

COMP	complementizer
CONJ	conjunction
DCOP	descriptive copula
dual	dual
excl	exclusive
F	factive
FOC	focus
HAB	habitual
IMPFT	imperfective
LOC	locative copula
N	noun
NEG	negation marker
NFAC	non-factive
OBLM	oblique marker
PERF	perfective
PL	plural
POSS	possessive
PROG	progressive
QM	question marker
RED	reduplicated verb root
REFL	reflexive
REL	relativizer
sg	singular
TOP	topic marker
V	verb
1,2,3	first, second, third person
* (asterisk)	indicates an ungrammatical sentence
: (colon)	indicates a lack of one-to-one correspondence between the Koh example and the English morphemes
- (hyphen)	indicates a one-to-one correspondence

References

Bamgbose, Ayo. 1974. On serial verbs and verbal status. The Journal of West African Languages. Vol IX no. I:17-47.

Boyd, Raymond. 1974. Étude comparative dans le groupe Adamawa. Bibliothèque de la SELAF 46. Paris: SELAF.

Comrie, Bernard. 1976. Aspect. Cambridge: Cambridge University Press.

1981. Language universals and linguistic typology. Chicago: University of Chicago Press.

Greenberg, Joseph H. 1963. Languages of Africa. The Hague: Mouton.

Johnson, Marion. 1980. A semantic description of temporal reference in the Kikuyu verb. Studies in African Linguistics, II(3):269-320.

Lyons, John. 1977 Semantics Volume 2. Cambridge: Cambridge University Press.

Samarin, William. 1971. Adamawa-Eastern. Current trends in linguistics, ed. T. Sebeok. 7:213-244.

Ubels, Edward. 1983. Mood and aspect in Karang. Yaoundé: SIL mimeo.

Ubels, Edward, and Virginia Ubels. 1980. Phonology of Karang. Yaoundé: SIL mimeo.

Wallace, Stephen. 1982. Figure and ground: the interrelationships of linguistic categories. Tense-aspect: between semantics and pragmatics. Paul J. Hopper, ed. 201-226. Amsterdam: John Benjamins Publishing Company.

Welmers, William E. 1973. African language structures. Berkeley: University of California Press.

RED LAKE FALLS, MINNESOTA: A SOCIOLINGUISTIC SURVEY

James Kapper

TABLE OF CONTENTS

Chapter I. Introduction....................................285
Chapter II. Methodology....................................288
Chapter III. The Community.................................293
Chapter IV. Survey Results.................................301
Chapter V. Conclusions.....................................315
Appendices...317
    Appendix A...317
    Appendix B...321
Bibliography...339

LIST OF TABLES

TABLE 1. Minnesota Migration Patterns before 1920.......296
TABLE 2. Red Lake Falls Immigrant Population............298
TABLE 3. Languages Spoken in Red Lake County............298
TABLE 4. Fluency in English of Persons Speaking a Language Other than English in the Home..............299
TABLE 5. Red Lake County Residents Claiming Single and Multiple Ancestry...............................300
TABLE 6. Age Distribution of Participants...............301
TABLE 7. Classification and Age Groups of Survey Sample...302
TABLE 8. Ancestry.......................................302
TABLE 9. Native Language Acquisition by Generations.....304
TABLE 10. Native Language Acquisition by Age Groups.....304
TABLE 11. Comparison of Native Languages by Age and Generational Groupings..............................305
TABLE 12. Language Usually Spoken in Past Situations, by Age Group..306
TABLE 13. Education Level by Age Group..................307
TABLE 14. Church Attendance among Catholics Surveyed....308
TABLE 15. Current Frequency of Use of French among Age Groups..309

TABLE 16. Use of French in Selected Situations by Those over Seventy-Nine Years Old....310
TABLE 17. Use of French in Selected Situations by Those Sixty to Seventy-Nine Years Old....310
TABLE 18. Use of French in Selected Situations by Those Forty to Fifty-Nine Years Old....311
TABLE 19. Linguistic Performance of Those over Seventy-Nine Years Old....312
TABLE 20. Linguistic Performance of Those Sixty to Seventy-Nine Years Old....312
TABLE 21. Linguistic Performance of Those Forty to Fifty-Nine Years Old....313
TABLE 22. Linguistic Performance of Those under Forty Years Old....313
TABLE 23. Code Switching by Age Groups....314

CHAPTER I

INTRODUCTION

Red Lake Falls is a small farming community located at the confluence of the Red Lake and Clearwater rivers, in the valley of the Red River of the North in northwestern Minnesota. The area was first settled in the nineteenth century by speakers of French who were later joined by other immigrant groups, principally Germans and Scandinavians.

Despite the varied ancestry of the area's population, and a time span of more than a century for the mechanism of the American "melting pot" to have done its work, a relatively large number of people in the area still use the French language in their homes and describe themselves and their communities as "French."

The background of the community and the survival of the language pose sociolinguistic questions concerning the status of the language in the community, the users of the language, and the situations in which it is used. The center of this study is a survey which will provide a basis for describing the linguistic community in terms of these sociolinguistic concerns.

Chapter I of this study is a discussion of the various phenomena likely to be covered in such a survey. A discussion of the methodology and development of the survey (which is reproduced in Appendix A) is contained in Chapter II. Chapter III is a brief description of Red Lake Falls, with notes on historical background. Data gathered by the survey are discussed in Chapter IV with conclusions and questions for further research presented in Chapter V. Appendix B contains a few Red Lake Falls French texts with translation and discussion.

The goal of studies like this one was discussed by Einar Haugen (1972a), who acknowledges descriptive linguistics as the basis of sociolinguistics while rejecting the traditional descriptivist's "monolithic" approach to language. Haugen emphasizes the significance of the relationship between the language and society, and defines the goal of sociolinguistics as being the description of the social and psychological situations of particular languages and the effects of these situations on languages and their users.

With this goal, sociolinguistic studies take on an interdisciplinary character, rooted as they are in descriptive linguistics yet embodying some of the concerns of sociology, psychology, history, ethnography, and other fields. This sociolinguistic study of Red Lake Falls reflects Haugen's concerns as he stated them in basic form in his nine "ecological" questions for the study of individual languages:

1. What is the classification of the language in relation to other languages?
2. Who are the users of the language?
3. What are its domains of use?
4. What internal varieties does the language show?
5. What concurrent languages are employed by its users?
6. What are its written traditions?
7. To what degree has its written form been standardized?
8. What kind of institutional support has it won?
9. What are the attitudes of its users towards the language in terms of intimacy and status, leading to personal identification?

The most important aspect of Haugen's ecological questions is not that each has been asked singly but that they are taken as being interrelated and descriptive of a system. As a study of a small French-speaking linguistic community existing within a larger English-speaking one, this study takes questions two and three, "Who are the users of the language?" and "What are its domains of use?" as its central issues. These have been expanded to include questions of change in patterns of language use. Who were the users of the language fifty or 100 years ago when the community began to develop? How have the domains of language use changed during this time? The changing influences in the factors mentioned in Haugen's other questions will also be considered in relation to these issues.

Joshua Fishman (1972a) has also emphasized the importance of non-linguistic elements in sociolinguistic studies. His point of view and methodology are basically sociological in nature, emphasizing the importance of describing the ways language reflects cultural regularities. Fishman (1972b) develops this further by dividing the sociology of language (the term he prefers to sociolinguistics [1972c]) into two categories: "descriptive sociology of language," which seeks to describe language usage norms for particular speech communities; and "dynamic sociology of language," which seeks to account for "differential changes in the social organization of language use and behavior toward language."

This study's concern with a particular speech community and the changes that have taken place within it places it within Fishman's "descriptive sociology of language" category. The survey in this study has been designed to indicate what social factors in the community of Red Lake Falls have influenced the language and its use. It is intended to provide a basis for a further investigation with a more descriptive, quantitative approach.

The most obvious questions faced by the study are related to bilingualism, language shift and maintenance, and phenomena related to language contact. Haugen (1972b, 1972c) discusses the importance of a linguistic description of these phenomena. However, this survey will place more emphasis on discovering the social and psychological attitudes they reflect than on describing their effects on the language. A thorough description of the languages in use at Red Lake Falls has not been undertaken.

Fishman (1972d) relates the study of individual bilingualism with societal bilingualism (diglossa). Diglossa is said to exist in societies which recognize two or more functionally different languages, each expressing a distinct set of behavioral attitudes and values. Both sets of values are accepted into the community and seen as culturally legitimate.

Diglossa exists with bilingualism where social roles are compartmentalized. Diglossa without bilingualism occurs where there are two separate speech communities due to political or class structures. Bilingualism without diglossa is a product of individual linguistic versatility.

The implications of individual bilingualism for the speech community as a whole are demonstrated in Susan Ervin-Tripp's (1973) psychologically oriented discussion of the phenomenon. She maintains that language learning occurs with social learning in both native language and second language acquisition. Ervin-Tripp shows how socialization takes place through identification with a model, leading to the adoption of a new role. This would imply that a study of individuals in a bilingual community would reveal correlations between the degree of fluency in the languages involved and certain social variables indicative of participation in the given cultures. Further, she shows that not only are the individuals' attitudes changed by learning a second language but that communities undergoing language shift, where widespread bilingualism is a necessary factor, would also be forced into cultural change as well. Social values reflected in the dominant language would be adopted into the culture by virtue of their being present in the "new" language. Surveys of language change, therefore, must incorporate some means of determining cultural changes as well.

Language shift and its relation to bilingualism in American immigrant communities are discussed by Anwar S. Dil and Timothy J. Curry (1981). They make the point that bilingualism need not lead to mother-tongue shift and that the causes of societal bilingualism and shift are not necessarily the same. Their study, based on demographic data from censuses, shows a strong correlation between greater linguistic and ethnic diversity among immigrant populations and a decreased resistance to the acquisition of English. Other figures show the economic advantages of learning English for first generation immigrants in an urban-industrial setting, as well as the advantage of their children's improved scholastic performance.

Fishman (1966) examines language loyalty and maintenance efforts by Franco-Americans in New England. The group studied by Fishman immigrated to the United States at the turn of the century; however, there have been French settlements in New England since much earlier times. In 1960 there were over a million native speakers of French in the U.S., mostly concentrated in New England. Fishman attributes the survival of their language to a network of interlocking religious, educational, cultural, and fraternal organizations, and continuing immigration from Canada. In the past fifty years education in French, the Franco-American press, and organizations which played an early role in preserving the French language have all lost influence. The present generation of Franco-Americans, according to Fishman, has begun abandoning many ethnic traditions as well as the French language, following a general feeling within the group that their dialect was inferior to standard French or Quebecoise.

The group studied by Fishman is similar to the Red Lake Falls group in terms of place of origin and time of immigration. The forebears of both communities arrived in the late nineteenth century from eastern Canada. Differences in settling--one group settled largely in industrial communities, the other on farms--and the amount of continuing immigration from Canada and other contact with French Canadians probably have caused differences in language and language use in the communities. Determining the extent of these linguistic differences is outside the immediate scope of this study. However, some of the sociolinguistic similarities between the Red Lake Falls community and other French communities will be explored here.

CHAPTER II

METHODOLOGY

Early research at Red Lake Falls consisted of a series of impromptu interviews exploring various topics including personal background, local and family history, and language habits and attitudes. Sometimes language data were recorded during these conversations, depending on the willingness of the person being interviewed. The purpose at this stage was to acquire familiarity with the dialect and to look for patterns in the habits of language use of individuals and within families.

As more information on the French language and people of Red Lake Falls was gathered, a need developed for a formal survey which could serve as an instrument for gathering and organizing sociolinguistic data. The survey which evolved from these interviews, and which is the center of this study, therefore has two purposes: 1) to provide data for characterizing the French community at Red Lake Falls in terms of present and past patterns of language use, and 2) to provide a foundation for further interviews in a more precise sociolinguistic study.

Instead of being a detailed statistical study, the survey is qualitative in its approach. This is due primarily to the preliminary nature of the project and the size of the population sample available for study. Red Lake Falls, the largest of the towns originally settled by the French in that part of Minnesota, was chosen as the location of the survey partly because of the concentration of the population. However, taking into account the limits of time, the financial resources available for conducting a survey, and the availability of persons willing to participate, the estimated number of subjects was thirty to fifty. Although this is a small sample, care was taken to make the selection reasonably representative in terms of the ages, amount of time in the community, and educational background of the people participating in the survey.

Sociolinguistic studies like Labov's (1972) and Wolfram's and Fasold's (1974) have emphasized the importance of empirical methods and quantitative, statistical studies. The major difference between their work and the present study is that those were detailed studies of larger populations examining the correlations between a wide range of social variables and a specific set of linguistic variables. This study is designed to provide preliminary information indicating what social and linguistic variables may exist within a speech community of a more limited size. The survey and interviews also provide information on participants' behavior and attitude towards their language, which will be useful in setting questions for a more detailed analysis of the community.

The methodology of sociolinguistic surveys aimed at describing language situations and measuring sociolinguistic changes is discussed by Stanley Lieberson (1981), who recommends several elementary procedures for improving the accuracy of such surveys. Several of his procedures were relevant to the preparation of this study:

1. The use of retrospective questions for obtaining age specific data
2. Cross tabulating the language characteristics of parents with those of their children
3. Consideration of migration into and out of the area being studied

Retrospective questions relating to language use are included in the survey for the purpose of contrasting current and past linguistic patterns, both within and between generations. These questions provide information which permits later use of statistical measures in analyzing the speech community.

Questions on the survey also allow for the cross tabulation of parents' and children's language characteristics. Lieberson (1981:269) says that "not only does the surveyor want to obtain a cross-tabulation between parents' and children's mother tongues but also additional characteristics which distinguish those bilinguals who pass on an acquired language from those who do not." A series of open-ended questions is included in section D of the survey, and these are followed up in later interviews for the purpose of discovering what those distinguishing characteristics or attitudes may be.

Cross-generational data are obtained in two ways: from a series of questions about the languages used by the parents and grandparents of those surveyed, and by interviewing the children and grandchildren (where possible) of older respondents.

In designing this survey, effort has been made to insure that historical factors influencing language change, such as migration patterns or the presence of other ethnic groups, will show up in the data through questions of genealogy and in more specific conversations during follow-up interviews. It is recognized that a survey occurring at one point in time cannot completely gauge the process or directly identify causes of language shift, but, as Lieberson has suggested, the problems associated with such surveys, particularly those which are based on a small sample of the population, can be minimized.

The survey (see Appendix A) is divided into four parts. Section A, "personal data," provides information on the individual's background in the community. This information on basic social factors such as age, sex, place of birth, family size, present residence, and education, when compared with the data on the language habits of individuals and families, may reveal patterns of language use for the community as a whole. Another factor believed to affect language use, affiliation with particular ethnic and religious organizations, is covered in section A, questions eight and nine. Question eight also provides an indirect way to measure the individual's involvement with his/her church.

The last question in section A is a means of determining the individual's identification with a particular ethnic group. Subjects are asked to select from a list those terms that describe which ethnic group(s) they belong to. This question is related to other questions on the survey which deal with attitudes toward language and other ethnic traditions.

Section B covers family history. Genealogical data in this section include places of birth of parents, grandparents, and great-grandparents, as well as specific information on other residences of the parents of the respondents. Questions eleven and seventeen request the source of the information and are intended to gauge the accuracy of the responses. The remaining questions in section B are intended to provide information on mother-tongue shift in individual families. Questions eighteen through twenty ask participants to list the languages spoken by their parents, grandparents, and great-grandparents. Again, the source for the information is provided as a check for accuracy.

Responses to the questions in section B show the history of language use and shift within families and, by extension, within the larger community. In addition, they show individuals' degree of familiarity with their family history and awareness of their linguistic heritage.

Patterns of migration in past generations, both within, and to and from the community, will also be indicated by this section.

Section C contains questions concerning the language-related behavior of individuals and family groups.

Questions related to language acquisition in this section require the participants to list their first and second languages, and approximate age and situation at the time of acquiring the second. These questions are important in follow-up interviews because individuals in the community usually have vivid memories of their motivation and other circumstances surrounding their acquisition of English. Bilinguals were also asked to rate their own competence in speaking, understanding, reading, and writing English and French on a five point-scale. There was no provision made for an objective test to rate the respondents' competence in any of these areas. More important to this survey is how speakers feel about their ability to communicate in English and French.

Questions twenty-two and twenty-four, relating to languages studied in school or used in connection with religious activities, are designed to discover evidence of institutional influences on language use. These questions were also followed up in later interviews to obtain more details and related information.

In the area of language use, question thirty-two asks participants to estimate the frequency of their use of the French language in different social contexts in order to provide data on current patterns of language use for the community. Retrospective questions are also included to show changes in community-wide patterns and in the habits of the individual over a period of time. Another specific question of language use concerns code switching. Respondents are asked to list situations when switching from one language to another occurs and to speculate on what the reasons for switching might be. The wording of the question is open-ended in order to gather as much information as possible. When practical, these questions are followed up by observation.

The respondent's perception of changes in his/her language habits is the subject of question twenty-eight. Like some previous questions, this one is meant to be a beginning point for discussion of how and why patterns of language use in the community have changed.

Questions thirty-one and thirty-two ask respondents to note how frequently they hear people in their families use French. The family is divided into age groupings, and answers are given in terms of a five-point scale. The questions are constructed to show the frequency of the use of French in communication between generations as well as within each generation.

Section D, the "attitude survey," is designed to provide clues to individuals' attitudes toward the French-Canadian culture and language at Red Lake Falls. The four open-ended questions cover a broad range of topics, and, while the responses may be assessed subjectively, they are more important as a "springboard" to a further, more detailed interview. In a sense, these questions act as a catalyst in the interview, with conversations on these topics stimulating the discussion of other memories and ideas along the same lines.

Other questions in the survey are built around a set of assumptions and hypotheses about language and language behavior in general, and about Red Lake Falls French in particular. In contrast, the open questions in section D will be more likely to reveal new information and to bring previously unnoticed or unforeseen questions to light.

The questionnaire was tested in a small pilot study in Grand Forks, North Dakota. Survey participants were volunteers from L'Assemblée Française de Grande Fourche, a French-Canadian ethnic organization in Grand Forks, North Dakota. The purpose of this pilot study was to test the questionnaire as a survey instrument. As a result minor changes were made in the survey format, and a few questions were altered. The genealogy questions in the family history section seemed to pose a major stumbling block for those completing the questionnaire. Despite the Grand Forks respondents' lack of information, the section was left in its original form and didn't pose as much of a problem in Red Lake Falls.

Two other questions did present problems in Red Lake Falls. Question four was interpreted by some people to apply only to siblings still living, whereas the intent of the question was to determine the total number of children in the respondent's family and the total number, living or dead, of older and younger siblings. Question twenty-eight is ambiguous. The question should read, "Is the language you usually speak now a different language from the one you usually spoke when you were younger?" These problems became apparent very early in the interview process and were corrected by the interviewer administering the survey during the course of each individual or group interview.

In the field two methods were used in administering the survey. On an individual basis the survey was treated as an interview, the questions being read from the form by the interviewer, who also recorded the responses. This was by far the best method because it provided the most complete information. Fewer questions were skipped, and if the subject of the interview digressed from the main point with additional relevant information, the interviewer could make adequate note of it. In general, people were willing to say a lot more about a subject than they were to write about it. An additional advantage of this method was the personal element added by the interviewer. Since subjects were talking to a person instead of writing for a nameless "someone" who had written a survey somewhere, they were more willing to provide details and to cooperate in recording language data.

A different method was used in administering the survey to groups. Survey forms were distributed, instructions and the two ambiguous questions were explained, and the respondents were then on their own in filling out the questionnaire. The interviewer checked each form for completeness and requested responses to skipped questions. A problem with administering the survey in this way was its length and the number of blanks to be filled in. This was intimidating to some people, and respondents in groups were more likely to skip questions. Factors influencing this may have been the age or level of education of survey participants.

Typically, follow-up questions were the same for both group interviews and individual. Additional information on second language acquisition and use was usually collected. Respondents often volunteered more information on family history and "the way things were" in the community and at home when they were growing up. With groups, conversations on those topics were usually spontaneous.

Language data were collected in a variety of ways. A few groups were told ahead of time that they should expect to speak French, and French conversation was encouraged during the course of the interviews and the following conversations. The tape recorder was left running with the consent of the participants who, in time, got used to its presence. Another approach that worked well was for the interviewer to play songs or jokes recorded in previous interviews. This was usually enough encouragement for the subject to donate something from his own repertoire, or to at least converse in French. Material recorded was transcribed from the tape and, in the case of stories and songs especially, checked for accuracy with a native speaker.

Interview subjects at Red Lake Falls were recruited through the use of social networks. Use of this system of referrals is discussed by Lesley Milroy (1980). A social network is best described as a group of "acquaintances," "friends," or "friends of friends." "It is a mechanism for exchanging goods and services, imposing obligations and conferring rights." Using the mechanism of the social network places the fieldworker in the position of being both an insider and an outsider. Each person interviewed is asked to recommend the interviewer to someone else. Thus, many more doors were opened, and in less time, than would otherwise have been the case. The first group interview at Red Lake Falls exemplifies the use of this technique. An older woman in the community agreed to cooperate in the survey and invited four of her friends to her house to participate in the survey and to pass the evening in conversation. The interview was successful, and each person present recommended six to ten additional subjects. People in this second group always agreed to an interview when the name of the person who made the recommendation was mentioned.

Almost all of the interviews at Red Lake Falls were arranged in this way. This accounts, in part, for the inclusion of relatively few people under the age of fifty in the sample; the older people who were interviewed had few younger people in their social networks.

When younger people were recommended, they were also more likely to decline to be interviewed on the basis that they were too busy or simply did not want to bother. Usually they expressed a concern that they really could not contribute to the survey because they did not speak French or know much about it. People in this category were more likely to agree to an interview if the recommendation had been made by an older relative or peer than they were when recommended by an acquaintance or friend of their parents or grandparents.

Labov (1972) also devotes some discussion to the methodological problems of observing language in its social context. He says the most systematic data for linguistic analysis are obtained when speakers are paying the least amount of attention to monitoring their own speech. The presence of one observer adds to the formality of a situation, and means must be found to direct the speaker's attention away from this constraint. One way to accomplish this is to observe speakers in a normal peer group over a period of time. Another is to involve the speaker in discussion of a topic which recreates strong emotions. Milroy (1980) also notes the importance of collecting examples of vernacular speech as a data base for linguistic analysis and notes that such data are more easily collected from self-recruited groups. These techniques were employed to a degree in recording additional linguistic data during interviews.

CHAPTER III

THE COMMUNITY

The area covered by the survey has as its center the city of Red Lake Falls and its immediate vicinity in western Red Lake County, including the rural area between Red Lake Falls and Huot, approximately five miles to the southwest.

French influence in this area can be viewed as occurring in two periods. The first began when French explorers first probed into the territory west of the Great Lakes in the late seventeenth century. By the 1730's trading posts had been established in northern Minnesota where French-Canadian fur traders had lived since as early as 1718 (Benoit 1975). These early traders were largely of French and Indian background, the semi-nomadic Métis.

The Métis tradition represents varying degrees of fusion of French-Canadian cultural and linguistic elements with those of the

Cree and Chippewa. Virgil Benoit describes the Métis way of life in the Red River Valley of the North:

> Traditionally in the Red River area, the Métis moved collectively in such a way as to participate in four annual geocultural activities: 1. they trapped and traded in the wooded areas and along waterways; 2. they wintered in communities along the Red River and to the west-north-west in North Dakota; 3. they hunted the buffalo on the prairies; and, 4. they cultivated small plots in subsistence-like fashion in village-like centers. Their manner of living depended on their cultural habits, social preferences, and geographical location (Benoit, in preparation, p. 14).

The Métis were important to the fur trade as both traders in pelts, and as provisioners, supplying meat and pemmican.

In 1844 a trading station was established at Pembina which became an important link between the Red River settlement (at the site of Winnipeg, Manitoba) and St. Paul. Métis traders established regular trade routes and moved goods via the Red River oxcart. Later settlements were often established along these trade routes, often at locations selected by the Métis as in the case of Red Lake Falls (Red Lake Bicentennial Committee 1976). Trade along these routes continued until the 1860s; but, during this time, political and economic changes were taking place which would effect major changes in the Métis way of life.

Although Minnesota was organized as a territory in 1849 and achieved statehood in 1858, large tracts of land in the north were not open for settlement, being in the possession of the Pembina and Red Lake bands of Chippewa, who valued the land as a hunting ground (Bjornson 1969). It was not until 1863 that the United States government was successful in pressuring the Chippewa into signing a treaty ceding to the United States an area approximately 180 miles long and 127 miles wide, at a cost to the U.S. of $510,000 (Red Lake Falls Bicentennial Committee 1976). This treaty opened approximately three million acres of land to settlement (Benoit 1975), and with the completion of initial surveys and the arrival of the railroad, the way was paved for a period of immigration and agricultural development which began early in the 1870s.

The economic changes concomitant with the arrival of large groups of permanent settlers and concurrent political changes in Minnesota and Canada caused the departure of many of the Métis from Minnesota. According to Benoit:

> In an effort to maintain their way of life, many Métis moved west to Saskatchewan with such leaders as Louis Riel. After 1885 when Riel was executed for leading forces of resistance against the Canadian government, many of his followers returned to Manitoba and North Dakota, especially to the Turtle Mountains area (Benoit, in preparation: 17).

How many Métis remained in Minnesota is undetermined although historical records show that many did, such as Pierre Bottineau.

What is now Red Lake County was part of the land ceded by the Cree and Chippewa in 1863 and was settled in the late 1860s and early 1870s. "The west end of the county was originally settled by French-Canadians, and the east end by both Scandinavians and French . . ." (Boughton 1976). The man given credit for founding Red Lake Falls is the Métis voyageur and scout Pierre Bottineau. Bottineau was also one of the original settlers of St. Paul, and founder of the city of Osseo, Minnesota (Red Lake Falls Bicentennial Committee 1976). Bottineau chose the site of Red Lake Falls; was instrumental in bringing groups of French-Canadian settlers from St. Paul, Canada, and the East; and served in local government for several years. Founded in 1876, Red Lake Falls was incorporated as a village in 1881, and as a city in 1898.

The group of permanent settlers who began to arrive in the 1860s includes French-Canadians from eastern Canada who represent a second wave of French influence distinct from that of the early Métis influence in Red Lake Falls and northern Minnesota. In fact, the Métis presence was still felt, though to a lesser degree, because of social pressures accompanying the political and economic changes taking place since the 1840s. "It is in this period of social turmoil that one finds the roots of prejudices favoring a system of sociocultural institutions and values of a European nature as opposed to the customs and traditions of Indian and Métis life along the Red River" (Benoit, in preparation: 13).

Census figures for this period show the rapid population growth and large influx of immigrants. Polk County, which included Red Lake and Pennington counties, in 1880, had a total population of 11,460. Of that total 5,697 persons (48.7%) were "native born," and 5,736 (50.3%) were "foreign born." The 1890 census shows a 163% increase in total population, or 30,892 resident, 15,501 (50.17%) of whom were "native born" and 14,691 (49.83%) "foreign born." No figures are available on the specific national origins of immigrants in these groups. One segment of the population wasn't growing. The 1880 census counted 163 "civilized Indians." This number had decreased to sixteen in 1890.

After the turn of the century, immigration had slowed considerably. Statistics on the number of immigrants arriving in a given year are not available for individual counties, but the 1920 census gives figures for the entire state (see Table 1). There is no reason to assume that this pattern applies evenly to the whole state, but it does serve to indicate a general trend.

TABLE 1

Minnesota Migration Patterns before 1920

Time of Immigration	Number of Male Immigrants	Number of Female Immigrants	% of Total 1920 Immigrant Population (486,795)
1918-19	1,652	1,594	.66
1916-17	2,542	2,050	.94
1911-15	28,119	16,614	9.18
1906-10	32,514	19,521	10.68
1901-05	32,058	20,218	10.73
Before 1900	165,642	121,373	58.96

Data for 8.81% of the total is not available.

(Bureau of the Census, 1920)

The first population figures for Red Lake County with its present boundaries come from the census of 1910. Red Lake County, in that year, had a total population of 6,564. Of these, 1,706 (26%) were "foreign born," 3,773 (57%) were "of foreign parentage," and 1,083 (17%) were "native born." Seven hundred and thirty-five persons (12% of the total 1910 population) were first generation French-Canadian immigrants, 1,269 (9.33%) had French-Canadian parents, 249 (3.79%) were born of French parents, and two were first generation French immigrants. The population of the city of Red Lake Falls at this time was 1,885.

The register of births for the city of Red Lake Falls gives some indication of the composition of the population in the community's early years. Unfortunately, the entire register was unavailable for examination at the time of this study. The first page of the register is reproduced in the Red Lake County Bicentennial Committee's history (1976) and records births during the first two years of the town's history, 1876 to 1878. Of the thirty-two children born, twenty-nine were described as "French." Two are listed as "three-fourths French" (one of these a grandchild of Pierre Bottineau), and one is "one-half French." Both parents of sixteen of the children were born in Canada, fourteen entries listed both parents as born in Minnesota, one had one parent from Canada and the other from Missouri, and one listed both parents as coming from Wisconsin. This listing gives the name LaBrie and describes the child as "one-half French". LaBrie is also the name of the couple from Canada and Missouri. Other families with common surnames have common places of origin as well.

What the census and township records show is that Red Lake Falls, in the beginning, was a predominantly French community whose first settlers were families already established in the United States for at least a generation. These original French-Canadian and Métis settlers were soon joined by a larger number of more recent immigrants from

Canada, and, shortly thereafter, Europeans. The largest number of immigrants arrived in the first twenty years of the town's existence although declining numbers of immigrants continued to arrive after the turn of the century. A further examination of the register of births and other township and church records is needed to provide more detail on the background of the early French settlers and the arrival of other immigrant groups.

Once they had arrived, immigrants did not always stay. For example, in the nearby community of Gentilly it is estimated that forty-nine families left the area between 1882 and 1889 (Benoit 1975). In this study, preliminary interviews with residents of Red Lake Falls on the subject of family history showed that many families had distant relatives in Washington State because relatives of their ancestors had moved on to that location. Furthermore, many French-Canadian immigrants had originally settled in New England before moving to Red Lake Falls.

Traditionally the Catholic church has played an important role in the lives of French-Canadians. In 1878 the parish of St. Joseph's was established in Red Lake Falls. Until that time French-speaking Catholics in the area had been served by a missionary priest. In 1883 the congregation split, and St. Mary's church was established. St. Joseph's was predominantly French and was served by French and Canadian religious orders until 1952 when the two congregations were joined. In 1929 membership in the parish was 250 families; in 1976 membership was 300 families (Red Lake Falls Bicentennial Committee 1976). The establishment of the Lutheran Church at Red Lake Falls in 1882, and the Presbyterian around 1890, probably coincides with increases in the populations of other immigrant groups.

After the initial period of settlement and growth Red Lake Falls' population has remained fairly stable. The smallest population for the city after this period was reported in 1930 as 1,386. Since then the population has steadily increased to its present estimated level of approximately 2,000. The economy of the area is still based on agriculture, with some light industry.

According to the 1980 census figures, the total population of Red Lake County is 5,471. Of this number, sixty-two persons (1.13% of the total population) are "foreign born" (see Table 2). All but seven of these are naturalized citizens. Thirty-nine of the sixty-two immigrants arrived sometime before 1950. More detail about the ages and specific years of immigration of these people is not available from census data.

TABLE 2

1980 Red Lake Falls Immigrant Population

Place of Birth	Foreign-Born Persons (62)
Germany	9
Italy	2
Sweden	4
Other European	13
USSR	3
Korea	2
Canada	29

(Bureau of the Census, 1980)

Information on the ancestry of residents of Red Lake County in 1980 is collected in Table 5. Of those reporting their ancestry, 3,304 claimed to belong to a single ancestry group while 1,922 claimed multiple ancestry groups. Ancestry was not specified by 245 persons.

Of the 5,004 persons over the age of five years counted in the 1980 census, 538 persons (9.83% of the total population) speak a language at home other than English. Specific information is shown in Table 3. No attempt is made in the census to ascertain the individuals' native language, degree of fluency in the other language, or frequency of its use. However, persons who speak a language other than English in their home were asked to rate their fluency in English (see Table 4).

TABLE 3

Languages Spoken in Red Lake County

Language Spoken at Home	Persons 5-17 Years (1,362)	Persons 18 Years and Older (3,642)
English Only	1,356	3,110
French	-	248
German	-	43
Italian	-	2
Polish	-	4
Spanish	-	6
Other	6	229

(Bureau of the Census, 1980)

TABLE 4

Fluency in English of Persons Speaking a Language Other than English in the Home

Speak English Very Well	408
Speak English Well	101
Speak English Not Well or Not at All	29
5-13 Years Old	0
14-17 Years Old	0
18-24 Years Old	2
25 Years and Older	27

(Bureau of the Census, 1980)

The background and present composition of the population of Red Lake County pose a series of questions relating to the Red Lake Falls dialect of French. In particular, the Métis presence early in the history of the area can be expected to have had some influence on the language. It remains to be seen if there is any connection between Michif, the language of many Métis, and Red Lake Falls French.

A similar question could be asked about dialect difference within the group of French-speaking settlers. Did the dialects of settlers coming from such diverse areas as Missouri, Maine, and Quebec have many differences, and if so, how did these factors interact, and what was the effect of this interaction on the language?

The composition and distribution of the immigrant population, in both the past and present, have implication for this study in the area of language attitudes, patterns of language use, and mother-tongue retention. Some important observations here are the concentration of the French population in western Red Lake County, the maintenance of a separate parish by Catholic French-Canadians, the number of persons in the present population who claim only French ancestry, and the number of people who use French in their homes. "French" is the largest single ancestry group claimed in the 1980 census. The largest multiple ancestry group is made up of those claiming "German and Others" (see Table 5). The fact that there are more persons claiming to be of pure French ancestry than pure German, and fewer persons claiming mixed French descent than German, in connection with the above historical observation, implies four possible hypotheses: 1) that no significant attitude or other social factor accounts for this distribution of the population; 2) that the French were separated geographically or socially from other ethnic groups; 3) that the smaller population of other ethnic groups forced them to intermarry more than the French; and 4) that people tend to make claims about their ancestry based on reasons relating to status.

TABLE 5

Red Lake County Residents Claiming Single and Multiple Ancestry

Single Ancestry Group		Multiple Ancestry Group	
French	1,001	French and Others	877
Norwegian	938		
German	847	German and Others	1,062
Other	233		
Irish	92	Irish and Others	294
Swedish	86		
Polish	33	Polish and Others	39
English	28	English and Others	161
Dutch	25		
Scottish	13		
Italian	5		
Russian	3		

(Bureau of the Census, 1980)

Some 1980 census data have a direct bearing on the discussion of mother-tongue retention at Red Lake Falls. Of the 538 persons who use another language in their homes (see Table 3), 248 (46%) use French. German speakers constitute the next largest group of forty-three (7.9%). The ratio of persons claiming any French ancestry to those claiming any German is .98/1. The ratio of those who use French to those who use German is 5.76/1. The assumption is made here, although the questions of ancestry and language use are unrelated in the census, that those who use French are of French descent, and those who use German are of German descent. In this small midwestern farming community it seems unlikely that a "foreign" language would be used in the home, so these languages were probably learned within the family. The questions are, therefore, "Why are so many more people likely to speak French than German?" or, "Why do more persons of French ancestry retain their language than persons in other ethnic groups?" The survey seeks to provide clues for identifying the mechanisms of support within the French-speaking community at Red Lake Falls, and the weaknesses in communication with surrounding communities that allow this smaller linguistic community to exist within a larger, English-speaking one.

CHAPTER IV

SURVEY RESULTS

The survey was conducted in June of 1984. Thirty-four people, twenty-three female and eleven male, completed the survey and participated in follow-up interviews. Survey participants ranged in age from eleven to ninety-two years (see Table 6). All survey participants currently live in the immediate vicinity of Red Lake Falls, either in town or within seven miles. Some participants have lived at the same address for as long as eighty-two years; others, usually those retired from farming, have moved into town in recent years. Two of the survey participants were born in Canada; the remaining thirty-two were born in Red Lake, Polk, and Pennington counties. All have lived most of their lives in the Red Lake Falls area.

TABLE 6

Age Distribution of Participants
(34 total)

Age Group	Number Surveyed
5-19 Yrs.	2
20-39 Yrs.	1
40-59 Yrs.	8
60-79 Yrs.	16
Over 79 Yrs.	7

For the purposes of this survey the population is divided into four categories: first generation, those persons of foreign birth; second generation, children of immigrants; third generation, grandchildren of immigrants; and fourth generation, great-grandchildren of immigrants. Table 7 shows the four classifications broken down by age group. Persons in the second generation category may have one parent born in the United States, as is the case for one-third of those in this category. Third and fourth generations have at least one-half of their grandparents or great-grandparents listed as "foreign-born." Both individuals in the first generation classification arrived in the Red Lake Falls area with their parents, one at the age of five, the other at eleven.

TABLE 7

Classification and Age Groups
of Survey Sample (33 total)

	Classification			
Age Group	1st Gen.	2nd Gen.	3rd Gen.	4th Gen.
5-19 Yrs.	-	-	-	2
20-39 Yrs.	-	-	-	1
40-59 Yrs.	-	1	4	2
60-79 Yrs.	1	8	7	-
over 79 Yrs.	1	6	-	-

Participants in the survey were asked to select from a list the ethnic terms they use to describe themselves (see Table 8). This question was included partially so the sample could be controlled, but also as an indication of the individual's identification with a particular ethnic group.

TABLE 8

Ancestry (34 total)

French	17
American	5
French-Canadian	2
American and French	7
American and French-Canadian	1
French and German	1
French, French-Canadian, and American	1

Terms were selected with about even frequency in all age groups and classifications with one notable exception. The five who selected only "American" were members of the third and fourth generation. In follow-up interviews these people indicated that they were aware of their French-Canadian ancestry. No one gave a specific reason for making the choice, or indicated a conscious decision not to be associated with any ethnic group. The question appears to be not only one of ancestry but of language use as well. Those who chose only "American" spoke French "very little" or "not at all." Furthermore, they indicated that, even though they may have some ability to speak or understand it, they do not make any attempt to use the language.

Those in the third and fourth generation who chose "French" or "French-Canadian" with or without "American" can either speak some French, and occasionally do, or have at least a few phrases, or in one subject's terms, "buzz words," which they use on occasion.

The self-reported data on linguistic performance of five people is not enough evidence to form the basis of generalizations which extend to the entire community; however, other observations and material gathered during interviews support the above suggestion that ethnic identification is linked to use of the language. Historically, the term "French," as used in the Red Lake Falls area, has not been limited to describing the people from France. It includes French-Canadians and their descendants (as in the register of births), in other words, people who could be expected to speak French. Furthermore, when an older person taking part in the study was asked to recommend someone who could tell stories or speak the language, the response was often something like, "Oh, _____ can help you. He's a real Frenchman." The informant did not mean the man had come from France, only that he had a good command of the language. The same people also referred, on occasion, to people who speak English as "English people" or "English," even when they were talking about a group of Scandinavians or Germans.

The implication of these observations is that the French language has a function in the community of identifying its members. How this relates to other social factors within the community, particularly those associated with the cultural identity of the group such as language maintenance or the borrowing of "foreign" cultural elements, should be established through more detailed study.

English is the predominant language at Red Lake Falls; however, twenty-four (70%) of the thirty-nine people surveyed listed French as their first language. Survey questions relating to patterns of language use and acquisition provide information for characterizing the French community as being a speech community that has undergone mother-tongue shift rather than a stable bilingual community.

Since mother-tongue shift, by definition, occurs between two generations, dramatic changes in patterns of language acquisition might be expected. Relatively rapid shifts have been documented in Finnish (Karttunen 1977) and Norwegian (Haugen 1972c) immigrant populations. Table 9, however, shows an apparently gradual shift. This gradual shift could be explained by the fact that immigration continued in this area for an extended period, and families have therefore been in the area for varying amounts of time. Tables 10 and 11 provide further information on this point.

TABLE 9

Native Language Acquisition by Generations

Generation	English as First Language	French as First Language
1st (2)	0	(2) 100%
2nd (15)	(1) 7%	(14) 93%
3rd (11)	(4) 36%	(7) 64%
4th (5)	(4) 80%	(1) 20%

TABLE 10

Native Language Acquisition by Age Groups

Age Group	English as First Language	French a First Language
Over 79 Yrs. (7)	0	(7) 100%
60-79 Yrs. (16)	(1) 6%	(15) 94%
50-59 Yrs. (2)	0	(2) 100%
40-49 Yrs. (6)	(6) 100%	0
Under 40 Yrs. (3)	(3) 100%	0

TABLE 11

Comparison of Native Languages by Age and Generational Groupings

Generation	% over Age 50	% Native French Speakers
1st Gen.	100%	100%
2nd Gen.	89%	89%
3rd Gen.	72%	64%
4th Gen.	20%	20%

(For purposes of this comparison, the 40-59 year category is divided.)

The tables show that no one in the survey under the age of fifty years had acquired French as a native language, and only one person over the age of fifty years had acquired English as a native language, a woman sixty-four years old, in the third generation. Although it seems very unlikely that a change would be so abrupt, and there may be other, older offspring of French-speaking parents who acquired English as a first language, information gained from an interview with the woman (subject #21) and her mother (subject #2) shows that her case may indeed be atypical for the Red Lake Falls community.

Subject #2's mother was born in Canada and raised near Red Lake Falls. Her father was born and raised in a community in southern Minnesota. He spoke French and English fluently, preferring English. When subject #2 started school, the family switched from the use of French in the home to English. Subject #2 recalls how difficult it was for her to learn English at this time, and for this reason, knowing that eventually her children would be required to learn English, she taught them, including subject #21, English as their first language. Another reason is revealed in her response to question number thirty, where she rates her ability to speak French as "very little."

Information on the past habits of language use of individuals gathered as a response to question twenty-seven, which asked survey participants to indicate the language usually used in certain situations, shows a pattern of bilingualism with English gradually becoming more and more predominant until a complete shift in mother tongue occurred approximately fifty to fifty-five years ago (see Table 12).

TABLE 12

Language Usually Spoken in Past Situations, by Age Group

Situation	Persons over 79 Years		Persons 60 to 79 Years		Persons 50 to 59 Years	
	French	English	French	English	French	English
With Friends	60%	40%	43%	57%	0%	100%
To Older Siblings in School	60%	40%	56%	44%	0%	100%
To Younger Siblings Not in School	100%	0%	60%	40%	0%	100%
To Relatives	100%	0%	86%	14%	50%	50%

All survey participants over fifty years old said French was the main language used in their homes when they were children, with the exception of subject #21, discussed above.

None of the people surveyed who are under the age of fifty listed French for any of the choices, even though many of them do have some knowledge of the language. All of the groups reported that English was the language they used in school with the exception of one Canadian-born woman who was educated in Montreal. In fact, most said that English was the only language they were allowed to use in school, both in class and during recesses.

In explaining their choice of language used with friends, people were often careful to explain that French usually was spoken "with French friends and English with English." One eighty-four-year-old man explains that he spoke English with his German friends. There were several German families in his area although "not many people associated with the Germans." The changes between the oldest and second oldest group in the number of people using English with friends may reflect four possible trends: a rising number of non-French speaking immigrants in the area and the attendant need for English as a lingua franca; a trend for people to associate more with other ethnic groups already present in the area; a greater acceptance of the English language by the French; or, the most likely, a combination of these factors. Distinguishing between these factors is not as easy as it would seem because, unfortunately, the amount of contact between ethnic groups in the area during the early period of the history of Red Lake County is not easily documented. Early census figures, for example, do not distinguish between various immigrant groups.

There is virtually no change between the two older groups in the number of people who used French with older siblings while they were growing up. People in both groups often remarked that they used French at home and English away from home with this group. On the other hand, a large change occurred in the use of English with younger siblings not yet in school, which, in connection with the increased amount of English being spoken with relatives by the second oldest group, probably implies a greater acceptance of English by the French as a group. Apparently the younger children in this situation learned French as their first language and were being given a "head start" on their English for school by older brothers and sisters.

In discussing their reasons for changing from French to English, survey participants usually mentioned two in particular: the increased number of people in the area who didn't speak French, and the gradually lessening influence of relatives who had insisted on the family's speaking French.

Sometimes the main reason a person spoke French was to be able to communicate with older relatives, or because conservative parents insisted that the French language be spoken at home. When the child left home, or the older relatives died, there was a decreased need, or opportunity, to speak French. The latter is the situation in the case of a man seventy-three years old who says his main language was French until seven years ago when his father died. The necessity of maintaining French for communication with older family members shows up in the fifty-fifty split between English and French in this category among the fifty to fifty-nine years old group (see Table 12). This is the only category other than "in the home" in which French was usually used by any members of this age group.

Education was also an important factor in the shift from the French mother tongue to English. The school setting provided the child's contact with the English language as well as contacts with children of other ethnic groups. Although this study has no direct way of measuring the relationship between fluency in French and English and the number of years spent in school, it seems likely that there is a correlation between the amount of education an individual had and the likelihood of his/her teaching English to his/her children as a first language. Table 13 shows the mean number of years in school for members of each age group.

TABLE 13

Education Level by Age Group

Age Group	Number Surveyed Who Attended College	Avg. Number of Years in School for Non-College
Over 79 Yrs.	3	6.6
60-79 Yrs.	0	7.8
40-59 Yrs.	2	11
Under 40 Yrs.	2	8*

*Non-college participants under 40 who are still in school.

The church has been proposed as a strong influence for language maintenance among the Franco-American communities of New England (Fishman 1966). Survey data and information from interviews seem to indicate a connection between the church and language maintenance, or shift, at Red Lake Falls. Whether changes in the local parishes influenced the situation in the community or only reflect it cannot be determined by this survey.

Of the thirty-four persons surveyed, thirty-two attend church. All thirty-two are Roman Catholic. Table 14 shows frequency of church attendance in order to demonstrate individual involvement with the church as a possible indirect indication of the church's potential as an influence or mirror of linguistic behavior.

TABLE 14

Church Attendance among Catholics Surveyed

Number of Individuals (32*)	Number of Masses Attended per Month
9	1-4
16	5-6
2	6-10
3	11-13
1	16-20

*One person, surveyed but not represented in these figures, was a priest.

None of the thirty-two Catholics surveyed reported the contemporary use of French in connection with religious services. Twenty-two of them reported that at one time French was regularly used at the church they attended. Several elaborated in interviews, saying that as recently as thirty years ago one of the Sunday masses at St. Joseph's, in Red Lake Falls, was in French. This would have been up to the time the parishes of St. Joseph's and St. Mary's were reunited. Another interview subject gives the year 1938 as being the last year French was used at St. Peter's church in Gentilly. This time period for the decreasing use of French in the mass coincides with the decreasing acquisition of French as a native tongue.

In order to describe current patterns of language use by individuals in the community, survey participants were asked to indicate how frequently French is used by various age groups in their families (see Table 15).

TABLE 15

Current Frequency of Use of French among Age Groups

Age Group	Frequency Always	Over 1/2 Time	1/2 Time	Less than 1/2 Time	Never
20-40 Yrs.				11%	89%
40-60 Yrs.				59%	41%
Over 60 Yrs.	3%	7%	19%	58%	13%
Between Those 40 to 60 and Those over 60			15%	63%	22%
Between Those 20-40 and Those over 60				8%	92%

(All other categories in question thirty-one were 100% "never.")

Question thirty-two of the survey asks individuals to report how frequently they use French in a variety of situations. People under forty years of age reported that they never use French in the situations listed. Tables 16, 17, and 18 summarize the responses of the remaining age groups. Not all of the people who reported speaking French in these situations are native French speakers. There was no apparent connection between a speaker's generation and his/her reported use of the French language.

TABLE 16

Use of French in Selected Situations by
Those over Seventy-Nine Years Old

Situation	Frequency				
	Always	Over 1/2 Time	1/2 Time	Less than 1/2 Time	Never
At Parties with Friends		14%		28%	58%
Elsewhere with Friends			28%	58%	14%
With Relatives		33%	17%	33%	17%
At Family Gatherings		40%		40%	20%

TABLE 17

Use of French in Selected Situations by
Those Sixty to Seventy-Nine Years Old

Situation	Frequency				
	Always	Over 1/2 Time	1/2 Time	Less than 1/2 Time	Never
At Parties with Friends		12 1/2%	12 1/2%	25%	50%
Elsewhere with Friends			13%	56%	31%
With Relatives		6%	25%	63%	6%
At Family Gatherings			20%	33%	47%

TABLE 18

Use of French in Selected Situations by Those Forty to Fifty-Nine Years Old

Situation	Always	Over 1/2 Time	1/2 Time	Less than 1/2 Time	Never
	Frequency				
At Parties with Friends				37%	63%
Elsewhere with Friends				37%	63%
With Relatives				63%	37%
At Family Gatherings				63%	37%

Those surveyed were also asked on the questionnaire, and in the interview, to list other situations when they might speak French. No other situations were named.

Survey participants were also asked to rate their competence in the use of French and English. This information examined by age groups (see Tables 19, 20, 21, and 22) has implications for the analysis of both current and past patterns of language use.

TABLE 19

Linguistic Performance of Those over Seventy-Nine Years Old

	Skill Level				
Function	Very Well	Well	Some	Very Little	Not At All
Understand French	86%	14%			
Speak French	72%	14%	14%		
Read French		16%	28%	28%	28%
Write French	16%			27%	57%
Understand English	86%	14%			
Speak English	71%	29%			
Read English	57%	43%			
Write English	57%	43%			

TABLE 20

Linguistic Performance of
Those Sixty to Seventy-Nine Years Old

	Skill Level				
Function	Very Well	Well	Some	Very Little	Not At All
Understand French	67%	23%	10%		
Speak French	56%	31%	13%		
Read French	7%		13%	40%	40%
Write French	7%	7%	7%	7%	72%
Understand English	75%	19%	6%		
Speak English	67%	23%	10%		
Read English	57%	37%		1%	5%
Write English	42%	44%		7%	7%

TABLE 21

Linguistic Performance of Those Forty to Fifty-Nine Years Old

Function	Skill Level Very Well	Well	Some	Very Little	Not At All
Understand French	38%	25%	12%	25%	
Speak French	11%		28%	50%	11%
Read French	13%		13%	13%	61%
Write French	12 1/2%			12 1/2%	75%
Understand English	88%		12%		
Speak English	88%	12%			
Read English	88%	12%			
Write English	88%		12%		

TABLE 22

Linguistic Performance of Those under Forty Years Old

Function	Very Well	Well	Some	Very Little	Not At All
Understand French			33 1/3%	33 1/3%	33 1/3%
Speak French			34%		66%
Read French			34%		66%
Write French					100%
Understand English	100%				
Speak English	100%				
Read English	100%				
Write English	100%				

Current patterns in the use of French in Red Lake Falls are a result of the interaction of various factors including the habits of French speakers who continue to use the language in the same situations they always have. For example, some people speak French with some friends and English with others--even when those others speak French--for no other reason than habit. Lesser social pressures, such as the need to be polite, are also factors. It is considered very rude to speak French in the presence of anyone who may not understand it. For these reasons it is possible for two French speakers to know each other for many years and for neither one to realize that the other speaks French. Other social and psychological factors may include the possible "group identification" functions of use of the language, and individual attitudes toward the status of the language. Many people feel that it is somehow inferior to European or Canadian French. Fluency also plays an important role in establishing social patterns. Even though a speaker may not be fluent, there are still times when he/she will use French.

Another aspect of the patterns of language behavior considered in this study is code switching. The older speakers in the survey, also the most fluent group, tended to code switch more than younger, less fluent groups (see Table 23).

TABLE 23

Code Switching by Age Groups

	Reported Switching	
Age Group	From French to English	From English to French
Over 79 Yrs.	100%	100%
69-79 Yrs.	81%	0%
40-59 Yrs.	33%	33%
Under 40 Yrs.	33%	0%

A limited amount of code switching was observed and the topic was discussed in interviews. Reasons the subjects gave to explain code switching are that some things "sound better" to them in one language or the other, and that "you can't always say what you want in French." There was no pattern in the topics people mentioned as being likely to cause them to change languages. Speakers who are fluent apparently code switch for effect, to make a joke in one language or to take advantage of certain nuances of meaning or sound. Speakers who are not fluent switch back to their native language to express their basic meaning. This accounts for the one-way switching in those under forty. In either case the speakers are making maximum use of the systems available to them.

CHAPTER V

CONCLUSIONS

Indications from the survey data are that the French community at Red Lake Falls existed as a relatively stable bilingual community from the time of its establishment in the early 1870s until the mid-1930s. During this sixty-year period, French-speaking parents continued to pass French on to their children as a mother tongue. Also during this period, the use of English was increasing in the community.

Factors responsible for the increasing tendency for French residents of the community to speak English include a decreasing number of new French immigrants, the presence of a relatively large population of non-francophone settlers, decreasing contact with the French language spoken in Canada and other areas of the United States, and, probably the largest single influence, public education in English. Since the establishment of Red Lake Falls the first language children were taught to read and write had been English. At the end of sixty years the French at Red Lake Falls were probably completely separated from the written traditions of their language, and from the support for language maintenance such traditions provide.

Institutional support from the Catholic Church was also removed at this time when French was no longer a part of religious services. Catechism in French was also replaced by the Baltimore Catechism in English.

The end result of the interaction of these factors was a seemingly abrupt period of mother tongue shift which took place in the 1930s. Although the survey data seem to indicate that there was a clear point in time when children of native French-speaking parents acquired English instead of French as a native language, use of French outside the home had apparently been on the decline for some time before this. By this time the only real reason to speak French was to communicate with older, monolingual relatives.

The indications of linguistic conservatism shown by the survey data for speakers of French, and their parents and grandparents, are supported by indirect evidence from census data which indicates that the French had a greater tendency to retain their language than other ethnic groups in the area. This is also a characteristic of French groups studied elsewhere in the United States (Fishman 1966). Other evidence in the survey implies that the language has an important function in establishing ethnic identity. This function of the language probably accounts for the linguistic conservatism at Red Lake Falls.

Responses to the attitude section of the survey indicate that people perceive virtually no difference between the French and other ethnic groups in the area. Older participants indicate that they have watched the differences disappear with time. This implies not the merging of diverse traditions but the disappearance of a cultural identity. The identifying traits of the French ethnic group mentioned by people responding to these questions were food and language. While many of the younger people surveyed proclaimed a dislike for pea soup and blood sausage, respondents to question number thirty-six were unanimous in saying they enjoyed hearing French spoken, even though they might not completely understand it. Sadly, the resources for preserving the language so that this will be possible beyond the near future are rapidly disappearing, the most fluent speakers of the language having reached their seventies and eighties.

In responding to question thirty-six, one seventy-six-year-old man said he was "proud to be French" and liked to hear the language, but strongly felt that "groups shouldn't try to be different from each other. You take the best parts from each group and use it to build something new, and that's what America did." Another man who also liked hearing the language used did not think it was a good idea to teach children French because "it mixes a kid up to learn both languages." Others, expressing a feeling of loss, said they regret not teaching their children French. Some have even tried, without much success, to teach their grandchildren the language, the difficulty, as they see it, being that they do not have enough contact with the children to be able to teach them more than a few words.

Concerning the linguistic description of Red Lake Falls French, only very broad generalizations can be made from the data gathered as part of the survey. The texts in Appendix B are perhaps more interesting as cultural artifacts than linguistic data, representing, as they do, somewhat formal speech. Phonological variations exist between speakers, and within individual speech styles, but study of a larger body of data is necessary to determine which, if any, of the variations are of social significance.

The general attitude of French speakers at Red Lake Falls toward their language is that it is somehow debased in relation to "the real French" (standard European French). People often comment that it is "mixed up with English" or "broken." These last comments may be references to English loan words or code switching. Very little switching was observed, but some vernacular speech was recorded in which English connectives (and, but) were inserted into French sentences.

Vernacular speech is difficult to elicit for several reasons. People are reluctant to speak French in the presence of someone who they think does not understand it. On the other hand, if the researcher speaks a French dialect other than the one of the area, people are also reluctant to talk, either out of a fear of not being understood or out of a fear of a negative reaction to their language. More time than was available for this study is needed to overcome the barrier that these reactions create for the elicitation of data.

Along with a larger body of linguistic data the study of a larger population in future research might provide data for accurate statistical assessment of the language situations discussed in Chapter IV. Including other ethnic groups in the survey as well would allow the comparison of patterns of language retention and use in diverse groups, and provide a more detailed assessment of the sociolinguistic environment of the area.

APPENDIX A

LANGUAGE SURVEY

A. PERSONAL DATA

1. Name:_________________________ Date:__/__/__

2. Age:_____ Male___
Female___

3. What is the highest grade or degree you completed in school?_____

4. How many older brothers do you have?___Sisters?___
How many younger brothers do you have?___Sisters?___

5. Residence:_________________________
____
How long have you lived at that address?________

6. Where were you born?_________________________________

7. Where else have you lived?_________________________
For how long (all together)?__________

8. Do you attend church? yes___no___
If yes, approximately how many times a month?___
Which church?____________________

9. Do you belong to any clubs or organizations? yes__no__
If yes, which ones?_________________________________

10. Which of the following terms would you use to describe yourself? (More than one may apply.)
American___French-Canadian___French___German___
Canadian___Franco-American___Métis___Norwegian___
Native American___Other (specify)_______________

B. FAMILY HISTORY

In section B answer as specifically as possible. If you aren't sure of a city, name the state/province or country, if possible. If you don't know the answer to a question leave it blank. Sources for information may be things like observation, assumption, word-of-mouth, etc.

11. Have you, or has anyone you know of, traced your "family tree?" yes___no___
If yes, approximately how far back?__________

12. Where was your mother born?__________________________

13. Where was your father born?__________________________

14. Did either of your parents live outside the Red Lake Falls area? yes___no___
If yes, where?____________________
For how long?_______________

15. Where were your grandparents born?
mother's father:____________________
mother's mother:____________________
father's father:____________________
father's mother:____________________

16. Where were your great-grandparents born?
mother's father's mother:____________________
mother's father's father:____________________
mother's mother's mother:____________________
mother's mother's father:____________________
father's father's mother:____________________
father's father's father:____________________
father's mother's mother:____________________
father's mother's father:____________________

17. What is your source for this information?___________

18. What language(s) did your parents speak?
mother's "first" language:__________________________
other (specify):__________________________

BEST COPY AVAILABLE

father's "first" language:___________________________

other (specify):___________________________

19. What language(s) did your grandparents speak?

mother's father's "first" language:_________________

other (specify):_________________

mother's mother's "first" language:_________________

other (specify):_________________

father's father's "first" language:_________________

other (specify):_________________

father's mother's "first" language:_________________

other (specify):_________________

20. What language(s) did your great-grandparents speak?

father's family:____________________

mother's family:____________________

21. What is your source for this information?____________

C. LANGUAGE HABITS

22. Have you ever studied a language other than English in school? yes___no___

If yes, what language?____________________

Why? ___required

___personal interest

___parents' idea

___already familiar with the language

___other (specify)______________________________

23. Are you able to:

	YES	NO
understand French	___	___
speak French	___	___
read French	___	___
write French	___	___

24. Is French used in the services at the church you attend? yes___no___only on special occasions (explain)_________

____

Was French spoken in the church you attended when you were younger? yes, regularly___no, never___on the same occasions as above___don't remember___

25. What was the first language you learned to speak as a child? English___French___other (specify)____________

If French is your first language, how old were you when you learned English?________Where did you learn to speak English?____________________

26. What was the first language you learned to read and write?

English___French___other (specify)______________none___

27. When you were growing up, what languages did you usually speak in the following situations? (check one)

	French	English	other (specify)
at home:	___	___	________________
at school:	___	___	________________
with friends:	___	___	________________
to older brothers and sisters in school:	___	___	________________
to younger brothers and sisters not in school:	___	___	________________
to relatives:	___	___	________________

28. Is the language you usually speak now different from the language you usually spoke when you were younger? yes___no___

If yes, approximately when did you switch?___________

Was there a reason for changing that you were aware of? (If so please explain briefly.)

____

29. Do you ever begin a conversation in French then switch to English? yes___no___

Switch to French from English? yes___no___

Is there a reason that you are aware of? (If so, please explain briefly.)

____

____

BEST COPY AVAILABLE

BEST COPY AVAILABLE

30. How well are you able to:

	very well	well	some	very little	not at all
understand French	___	___	___	___	___
speak French	___	___	___	___	___
read French	___	___	___	___	___
write French	___	___	___	___	___
understand English	___	___	___	___	___
speak English	___	___	___	___	___
read English	___	___	___	___	___
write English	___	___	___	___	___

31. If a language other than English is known in your family what is the language?______________________

How frequently is it used in the following groups?

	always	over ½ the time	½ time	less than ½ time	never
among children (5-20)	___	___	___	___	___
among young adults (20-40)	___	___	___	___	___
among middle aged adults (40-60)	___	___	___	___	___
among elderly (over 60)	___	___	___	___	___
between elderly and middle aged	___	___	___	___	___
between elderly and young adult	___	___	___	___	___
between elderly and children	___	___	___	___	___
between middle aged and young adults	___	___	___	___	___
between middle aged and children	___	___	___	___	___
between young adults and children	___	___	___	___	___

32. How often do you use French in the following situations?

	always	over ½ the time	½ time	less than ½ time	never
at parties with friends	___	___	___	___	___
elsewhere with friends	___	___	___	___	___
with relatives	___	___	___	___	___
at family gatherings	___	___	___	___	___
other (specify) __________	___	___	___	___	___

33. In your family do people ever talk in French to other people who answer them in English? yes___no___

A. Some older people speak French to some younger people who answer in English:

always___usually___½ time___seldom___never___

B. Some younger people speak French to some older people who answer them in English:

always___usually___½ time___seldom___never___

Has either of these situations ever existed in your family? yes___no___

If yes, was it situation A or B?_____

D. ATTITUDE SURVEY

If your family is of French or Canadian origin please take a few minutes to briefly answer these questions. Use the back side of these pages if you need more space.

34. What are the things in your family that are 'typically' French-Canadian?

35. Are French-Canadians different from other nationalities in the area? How?

36. Do you like to hear French (even though you might not understand it)? Would you like to study French yourself, or have your children or grandchildren learn it? Why, or why not?

37. Who speaks (spoke) the better French in your family, men, women, neither? What members of the family still speak French? Did you speak French more with your mother (grandmother) or your father (grandfather)?

BEST COPY AVAILABLE

APPENDIX B

The stories and songs in this appendix were recorded and later transcribed using standard French orthography with some slight modifications made to accommodate the most obvious phonological variations found in Red Lake Falls French. In the case of Text #1 the international phonetic alphabet was used. When possible, transcriptions were verified with the source, or in some cases, with other native speakers. These transcriptions are presented here with English translations.

Red Lake Falls French is not usually written. Those few people who indicated in the survey that they have some knowledge of written French, rarely, if ever, use that skill, although a few of the older residents occasionally correspond with relatives in Canada. No examples of written forms were available for study. Because of the above factors, and the situation of Red Lake Falls French in relation to English, there is no standard orthography which accurately represents the language.

To apply the standards of the orthography of Parisian French to this unwritten dialect, which has been probably separated from standard European French since the seventeenth century, is misleading though in some ways useful. First, it places the reader in familiar territory, so to speak; and, second, until the phonological and syntactic structures of Red Lake Falls French are more completely understood, looking at the dialect in this way provides points for comparison and contrast useful for describing the language.

In effect, with the texts transcribed in what is more or less standard French orthography, Red Lake Falls French has been represented as if the phonological component or the language were European French. (This may be checked against the phonetic transcription in the first text.) The sentence structure, however, is that of Red Lake Falls. Colloquialisms and items of Canadian or apparently local origin have been retained. Some of these points of usage may be unknown or ungrammatical in Parisian French.

The first text presented in this collection is a personal experience story related by a woman eighty-four years old. The story is typical of its kind, which most of the French speakers in Red Lake Falls can tell. This type of story is also the kind most often offered in an interview.

These personal stories have a polished sound to them, a degree of formality in their structure and the manner in which they are delivered which indicates that they have been often repeated. Some interview subjects have explained that these are the only kinds of stories they tell. There seems to be a general feeling in the community, among older residents especially, that it is important to be able to tell an entertaining story of some kind. Elements common in this kind of story are also basic to other kinds as well; unusual or humorous events, circumstances revealing or arising because of some character trait in an individual, and, a well timed and well deserved sarcastic comment for a "punchline." The characteristic that sets the stories apart from others is their personal nature. Although they often involve experiences common to the background of all the members of the community, they are related by each individual in his/her own way and with his/her own insights.

Text #1

1. a lãnʒ də sæŋk ɔn ʒete pardzy dãən bwoo
A l'age de cinq ans j'étais perdue dans le bois.
At the age of five I was lost in a woods.

2. ʒo vulja swiv ma suər por æle a lɛkɷl
Je voulais suivre ma soeur pour aller à l'école.
I wanted to follow my sister in order to go to school.

3. mee ʒe tɾuve pardzy mon pɛɾ e mameɾ
Mais je m'ai trouvée perdue. Mon père et ma mère,
But I found myself lost. My father and mother

4. e mɔ̃ tɾuve taron apre midzi ʒete fætsige
ils m'ont trouvée tard dans l'après-midi. J'étais fatiguée
found me later in the afternoon. I was tired

5. mee ʒe bɾooje ʒo məvwija ale alɛkɷl dãndəmæn
mais j'ai braillé; je voulais aller à l'école. Le lendmain
but I cried; I wanted to go to school. The following

6. matæ̃n mon pɛɾ ijdzi a mamɛɾ hæbij lã
matin, mon père il dit à ma mère, "Habille-la;
morning, my father said to my mother, "Dress her;

7. ɔnwijlã a lɛkɷl bjɛ̃n ʒete a lɛkɷl
envoie-la à l'école." Bien, j'étais à l'école.
send her to school." So I went to school.

8. mee la matɾɛs vwila aməfeɾ asir ɷn ʃɛz
Mais la maîtresse a voulu à me faire asseoir sur une chaise
But the teacher wanted to make me sit in a chair

9. eɾɛsee tɾankɩl mee mwe kəmdɾezon ʒɷneta pɔ
et rester tranquille. Mais moi, comme de raison, je n'étais pas
and be still. But I, of course, wasn't

10. kəpab dəɾɛse tɾankɩl eʒe kurəoje dɑ̃n lɛkɷl
capable de rester tranquille. Et j'ai courroyé dans l'école.
able to sit still. And I ran around in the school.

11. la matɾɛs adzi bæ̃ dɷmæn matæn rɛstʃevuu
La maîtresse a dit, "Ben! Demain matin restez chez-vous
The teacher said, "So! Tomorrow morning stay home

12. sitsuvɷ kuɾəoje ɔ̃ɔ̃ mee mwe ʒe bɾooje pi
si tu veux courroyer." Ah, mais moi j'ai braillé puis,
if you want to run around." Oh how I bawled because

13. ʒɷ vuɛzejale alɛkɷl landmæn ʒɷ vaa alɛkɷl
je voulais y aller à l'école. Le lendemain je vais à l'école.
I wanted to go to school. The next day I went to school.

14. la matɾɛs adzi bjɛ̃ɛ̃ tuɛɾvnʏu
La maîtresse a dit, "Bien, tu es revenue.
The teacher said, "Oh, so you came back.

15. fotsumekɷt kəm kɔstɛkɷtta meɾ
Faut que tu m'écoutes comme c'est tu écoutes ta mère."
You have to listen to me like you listen to your mother."

16. ʒɷ vɷɷkɷtstuɛsez asɷɷʃez la
"Je veux que tu t'assises dans cette chaise-là."
"I want you to sit in that chair there."

17. bjɛ̃ ɾɛstla tɷtla ʒuurne pi ʒe dzi bæ̃n
"Bien, reste-là toute la journée." Pis, j'ai dit, "Bien,
"Stay there all day." Then I said, "But,

18. ʒɷ nepa kəpabl arɛse asi tɾankɩl tɷɷt laʒuurne
je n'ai pas capable rester assise tranquille toute la journée."
I can't sit so still all day."

19. ʒɷ foka ʒɷ kuɾon wi bjɛ̃ɛ̃ ɛ adzi se koɾe
"Je faut que je courroie." "Bien," elle dit, "c'est correcte,
"I have to run around." "Well," she said, "that's true,

20. vɔ dahoɾ kuɾon tɔ̃ntsu vɷ me ətsurvjɛn
va dehors. Courroie tout que tu veux, mais que tu reviennes
go outside. Run around all you want, but when you come back

21. asitwe mɔdanedlvuraʒ afeeɾ
assieds-toi. M'en vais te donner de l'ouvrage à faire."
sit down. I'm going to give you some work to do."

22. ebjɛ̃ɛ̃ ɛamɷdənn dəkɾion dəkuluər
Eh bien, elle m'a donné des crayons de couleur,
So then she gave me some crayons,

23. æ̃n tsɪpapije ebjɛ̃n ʒepɔ̃sɛn laresɔn
un petit papier, eh bien, j'ai passé le restant
a little paper, and I spent the rest

24. də æpɾe midzi a ekɾiʀ əvɛk se kɾion
de l'après-midi à écrire avec ces crayons.
of the afternoon writing with the crayons.

25. pi ʒedzi bæ̃n maɾsi ʒɷ mɔnva ʃenuu
Pis, j'ai dit, "Bien, merci, je m'en vais chez nous,
Then I said, "Thank you, I'm going home now,

26. mee ʒnɷ vjɛndɾepɔ ebjɛ̃n ʒe pɔ ʀtɔrne
mais je ne reviendrai pas." Eh bien, je n'ai pas retournée.
but I'm not coming back." And I never went back.

Texts two through four are examples of a different kind of story told in the community. These are more a part of local lore and are more widely known in the community than the personal stories. People are more likely to tell the same ones in the same way. More like jokes than stories, these examples also have the same basic elements mentioned in the discussion of personal stories. Some examples contain some of the conventions of the fairytales and other traditional stories told in other North American French communities; for example, the character "Petit Jean" in text number four; or the opening line of number three, "Il avait un jeune couple là pis c'était ben pauvre." (Compared with the openings of numerous stories in the collection of French folktales from Missouri by Thomas and Thomas (1981) "C'est bon d'vous dzire, eune fouès, c'étaient ein vieux pis eune vielle . . . pis i'sontaient pauv's. . . .") Despite these apparent similarities, no one was found among those participating in the survey who said they remembered hearing "fairytales" or "ghost stories" of any kind.

Texts two and three are from the same source, a woman sixty-seven years old. The joke in number four is known in Red Lake Falls but was related by an individual who learned French outside the community and may not be a typical example of Red Lake Falls French.

Text #2

1. On était pour avoir de la visite du Canada
Some people were going to have company from Canada

2. et pis quand la visite est arrivée il
and when the company came they

3. était supposé amener leur grandmère.
were supposed to bring along their grandmother.

4. Mais, la grandmère elle n'a pas venue et pis
But, the grandmother didn't come and when

5. quand ils sont arrivés les gens ont dit, "Mais
they arrived the people said, "But

6. devous que est la grandmère?"
where is your grandmother?"

7. La grandmère a achetée..."
"Grandmother bought..."

8. "Quoi? La grandmère a achetée?"
"What? Your grandmother bought one?" [had a baby]

9. "Ben oui! M'en a achetée un capot de poil
"Yes! She bought a fur cloak

10. pour venir avec nous autre, et puis, avant
to come with us, then, before

11. de partir elle a tombée malade."
it was time to leave she got sick."

12. Ben il dit, "Elle va tu vivre?"
So he said, "Is she going to live?"

13. "Ah oui!" il dit. "Elle va vivre parce qu'elle est
"Oh yes!" he said. "She'll live because she's

14. bien trop mauvaise pour mourir."
too mean to die."

Text #3

1. Il avait un jeune couple là pis c'était ben pauvre
There was a young couple and they were very poor

2. parce-que le vieux était trop paresseux pour travailler.
because the husband was too lazy to work.

3. Et pis quand la vieille demandait dequoi il disait toujours,
And every time the wife asked for something he always said,

4. "Demain matin, on ira demain matin."
"Tomorrow morning, we'll get it tomorrow morning."

5. "Ben, il faut de grocerie là mon vieux." "Ah ben,
"We have to go to the grocery." "Oh well,

6. on ira demain matin." C'était toujours demain matin.
we'll get it tomorrow morning." Always tomorrow morning.

7. Ça fait qu'un soir ils sont couchés pis elle dit aux enfants
So, one evening, when they were in bed, she told the children,

8. "Demain matin on ne se lèvera pas. On n'a rien à manger
"Tomorrow don't get up. There's nothing to eat

9. dans la maison." On n'est pas pour se lever. On va rester
in the house." So they didn't get up. They stayed

10. couché. Pis le vieux s'est levé la vieille faisait à croire
in bed. When the husband got up, the wife pretended

11. qu'elle dormait. Pis le vieux descendu en bas. Pis, au bout
to sleep. The man went downstairs. After

12. d'une escousse il a dit, "Lève-toi la vielle, lève-toi." Pis,
a while he said, "Get up wife! Get up!" But

13. elle a pas répondu. D'un coup il crie encore, "Lève-toi,
she didn't answer. Shortly he cried again, "Get up

14. la vielle." "Pourquoi? "Ben," il dit, "C'est le temps de
wife!" "Why?" "Because it's time to

15. déjeuner." Pis elle dit aux enfants, "On va descendue là.
eat." She said to the children, "Go on down.

16. On va tout assoir à table en pensant que le vieux a fait
Sit at the table. Pretend he's made

17. le déjeuner." Pis ils sont descendus. Pis ils sont assis
lunch." So they went down and they sat

18. à table. Pis le vieux a dit, "Mais qu'est-ce que
at the table. When their father came down he said "What

19. vous faissez là?" Elle dit, "On attend pour notre déjeuner."
are you doing there?" The woman said, "Waiting for our lunch."

Text #4

1. Le curé était en train de visiter avec
The curie was making a visit with

2. la mère du petit Johnny. Et puis, le
little Johnny's mother. And so,

3. petit John, il est arrivé dans la maison.
little John, he came into the house.

4. Puis ils sont en plein hiver il font
Since it was full winter it was

5. ben frette. Et puis, naturellement, il
very cold. And, naturally, he

6. était dehors et il s'était habillé
was outdoors so he was dressed

7. avec deux paires de coulottes. Quand il
in two pairs of pants. When he

8. arrive dans la maison il a fait comme
came into the house he did as he

9. du coutume; il a ôté sa première paire
usually did; he took off his first pair

10. de coulottes. Il l'a posée sur les clous
of pants. He hung them on a nail

11. près de la porte. Et puis, pendu aux
next to the door. Hanging on

12. mêmes clous était une petite statue
the same nail was a small statue

13. du petit Jésus. Ça fait que quand
of little Jesus. It happened that when

14. il a ôté sa paire de coulottes, il a
he hung up his pair of pants, he

15. caché la statue du petit Jésus, et la
hid the statue of little Jesus. The

16. mère a parlé au curé et puis a voulu,
mother was talking to the curie and wanted

17. montrer que son fils était bien sage.
to show what a model son she had.

18. Et elle a dit à Johnny, "Ote tes coulottes
And she said to Johnny, "Take down your pants

19. de là pour que le curé puisse voir
over there so that the curie can see

20. ton petit Jésus."
your little Jesus."

The songs in texts five through seven are from the same source, an eighty-year-old woman who said she learned them from her mother. A version of number seven is still sung in Canada.

Number eight was recorded by an eighty-four-year old man. The song poses problems for translation because of some possibly archaic references (for example cous-cous d'un loup, pistoles) and because of its numerous plays on words and use of nonsense. (How much of the latter was in the original song and how much came to be added inadvertently as parts of the song were forgotten and changed over the years is unknown.) The man who recorded the song, in explaining two of the puns, says that planche de tanche is a carpenter's term for "tongue and groove" construction, and that pain d'épice tout garnis de saucisses refers to uncured boards of knotted pine covered with sap.

The song fragment in text nine is from a seventy-year-old woman who says this is part of a song her father used to sing. She doesn't know what language, other than French, the song is composed in, or the meaning of the non-French words, although she feels certain that "it's an Indian song."

Speakers of Michif and Ojibway have failed to recognize any of the non-French words in the song; however, a Lakhota speaker from the Standing Rock Reservation in South Dakota has pointed out several similarities between the language in the final part of the song and her Lakhota dialect. [tumɩnigatee] is similar to temni meaning "to sweat," and also mni, "water."* A more striking similarity is the word [wɩʃtee], which is like waśte in both its sound and distribution. Waśte is "good" or, as a response to a statement, "I agree" or "you know."

Text #5

1. C'était la jeune Rosalie qu'il s'en allait à la confesse.
   It happened that young Rosalie went to make her confession.

2. Mais son amant qui la voyait venir y joue t'un tour,
   But her lover, who saw her coming, played a joke,

3. par adresse. (Bien frontement y'a mis d'un surplis
   with skill. (With daring, he put on a surplis

4. et dans le confessional il s'est mis.) [repeats]
   and let himself into the confessional.)

5. Mon père je suis ici devant vous avec un
   Father, I am here before you with a

6. air bien répondante, je me confesse à vos
   very repentent feeling, I confess myself at your

7. genoux d'avoir été longtemps absente, (j'ai
knees to have been long absent [from confession], I have

8. parlé mal contre mon prochain et j'ai eu
spoken evil against my neighbor and I have

9. tort avec des saints.) [repeats]
wronged myself with the saints.

10. Mon père je m'accuse d'avoir péché, j'ai
Father, I accuse myself of sinning, I have

11. souvent fait la paresseuse, mais par avant
often been a gossip, and, before

12. de repartir j'ai t'un petit mot à vous dire
leaving now I have one small thing to tell you.

13. (Si vous voulez véritablement m'écouter un
If you really want to, listen to me

14. petit moment.) [repeats]
a moment.

15. Mon père j'ai t'un petit cavalier, qu'il me rend
Father I have a lover who has made me

16. vraiment amoureuse. Il m'entretien z'entre
truly love him. He holds me in his

17. ses bras, des mots d'amour, mais par adresse.
arms, his words are of love, but cunning.

18. (Il a goûté z'à mon malheur qui on a
(He has relished the unhappiness which

19. coûté cher à mon coeur.) [repeats]
has greatly troubled my heart.)

20. Ma belle retournez-vous en paix car'z
My lovely, go in peace because

21. vous vivrez toujours de même. (Car c'est
you always will live the same way. (Because it is

22. moi-même qu'il est ici, voudriez vous me
I, that one, who is here. Will you

23. prendre pour marie?) [repeats]
take me as a husband?)

24. Messieurs [je] dis vous véritablement,
Sirs, I tell you truly,

25. c'est z'y vous même qui me confesse.
It's you here who confess me.

26. Je m'accusais d'avoir tout dit.
I acted as if I told it all.

27. Grand Dieu d'honneur j'ai eu l'adresse
Great God of honor! I had the cleverness

28. (d'avoir caché un petit péché, peut-être
(to have hidden a small sin or perhaps

29. qu'il aurait tout gâté.) [repeats]
it would have spoiled everything!

Text #6

1. En me revenant de France j'ai passé par
Coming back from France I passed through

2. St. Malo. J'étais arrêté dans une barse
St. Malo. I stopped in a tavern

3. pour y boire du vin nouveau.
to drink some new wine.

4. (refrain) (C'est à boire à boire mesdames, c'est
Oh to drink, to drink, my ladies,

5. à boire qu'il nous faut.)
we must drink.

6. J'étais arrêté dans une barse pour y boire
I stopped in a tavern to drink

7. du vin nouveau. J'ai en demandé t'une
some new wine. I ordered a

8. chapine, elle me donne z'un rire t'un pot.
pint, she gave me the tankard and smiled.

(refrain)

9. J'ai en demandé t'une chapine, elle me
I ordered a pint, she

10. donne z'un rire t'un pot, et quand c'a
gave me the tankard and smiled, and when it

11. venu pour payer elle m'a fait payer la peau.
came time to pay she made me pay with my skin.

(refrain)

12. Quand c'est venu pour payer elle m'a fait
When it came time to pay she made me

13. payer la peau. Elle m'a fait vendre mes
pay with my skin. She made me sell my

14. coulottes, ma casquette et mon manteau.
pants, my hat, and my coat.

(refrain)

15. Elle m'a fait vendre mes coulottes, ma
She made me sell my pants, my

16. casquette, et mon manteau. Et quand c'a
hat, and my coat. And when the time

17. venu pour partir elle m'a fait monter en haut.
came to leave she made me go upstairs.

(refrain)

18. Quand c'est venu pour partir elle m'a
When the time came to leave she

19. fait monter en haut, et la bonne femme
made me go upstairs, and the old crone

20. qui est par derrière qu'il dit "cache ton
who was in the back said "[time to go] hide your

petit oiseau."
little bird."

(refrain)

Text #7

1. I took the road mon petit panier sur mon bras.
with my little basket on my arm

2. On my road I met le fils d'un avocat.
a lawyer's son

3. (refrain) Oh Dear I love you ah oui à ma manière
oh yes, in my own way

4. Oh Dear I love you oh non vous m'aimez pas.
oh no, you don't love me

5. On my road I met le fils d'un avocat.
a lawyer's son

6. He asked me what I had dedans ce panier-là
in the basket there

(refrain)

7. He asked me what I had dedans ce panier-là
in the basket there

8. I told him it was some apples "mais vous n'en vouliez y pas."
but you wouldn't want any.

(refrain)

9. I told him it was some apples "mais vous n'en vouliez y pas."
but you wouldn't want any

10. "I'll take a dozen mon père vous les paiera."
my father will pay you

(refrain)

11. "I'll take a dozen mon père vous les paiera."
my father will pay you

12. Go into the other room le bon homme il est là.
the old man is in there

(refrain)

13. I went into the other room le bon homme il n'y était pas.
the old man wasn't in there

14. So I went upstairs le bon homme il était là.
the old man was there

(refrain)

15. I went upstairs le bon homme il était là.
the old man was there

16. I asked him for my money il m'a dit qu il n'avait pas.
he told me he didn't have any

(refrain)

17. I asked him for my money il m'a dit qu'il n'avait pas.
he told me he didn't have any

18. I told him to go to hell le maudit avocat.
the damned lawyer.

(refrain)

19. I told him to go to hell le maudit avocat.
the damned lawyer

20. I'll sell some more apples mais vous en n'auriez pas
but you won't have any of them.

(refrain)

Text #8

1. Dans le pays de l'Angleterre
In the country of England

2. Dans la vieille d'où je viens
In the old times, where I come from

3. J'ai vu dans le désert
I saw, in a clearing

4. Des cou-cous d'un loup
Some naughty wolves

5. Tout habillés en forme de bataille
All dressed-up and in battle form

6. Des rouges et pis des coyes
In red and spots.

7. Que ce pays est drôle!
What a funny country that is!

8. Tout du long du bois
All along the woods

9. La bouteille a du vin
The bottle has some wine.

10. Et dedans ce pays-là
And in that country,

11. Que ce pays-là est drôle,
What a funny country!

12. On dort à profond, someille
We sleep soundly

13. sans que personne nous réveille
Without anyone waking us,

14. Et ne rien que chanter
And doing nothing but singing

15. On gagne de bel argent
We make good money

16. Et à jouer on gagne des "pistoles"
And win pistoles gambling

17. Que ce pays est drôle!
What a funny country!

18. Tout du long du bois
All along the woods

19. la bouteille a du vin
The bottle has some wine.

20. Et dedans ce pays-là
And in that country

21. Que ce pays-là est drôle!
What a funny country!

22. Les maisons sont faites de planche de tanche
The houses are made of fish

23. Pour les ceux qui veulent en prendre
For those who want to take some

24. Les couvertures de jambon
The roofing of ham

25. Les murs de pain d'epice
The walls of gingerbread

26. Tout garnis de saucisses
All covered with sausages

27. Que ce pays est drôle!
What a funny country!

28. Tout du long du bois
All along the woods

29. La bouteille a de vin
The bottle has some wine.

Text #9

1. Il avait un petit sauvage
There was an Indian boy

2. tout noir tout barbouillé wɪʃtʰee
all black and dirty-faced

3. Il était à la rivière
He went to the river

4. c'était pour se débarbouiller wɪʃtʰee
to wash his face

5. /tuumɪnigatee watʰee wɪʃtʰee

6. /mɪnigee watee watee mɪnige wɪʃtʰee/

NOTES

*In the Oglala dialect *m.nɩt'a'* is "drown" as in *wa'ta wa m.nɩt'a'*. "ship/a/it drowns" (Beckwith 1930: 420).

References

Agheyisi, Rebecca, and Joshua A. Fishman. 1970. Language attitude surveys. Anthropological Linguistics, 12: 137-157.

Barbeau, Victor. 1970. Le français du Canada. Quebec: Librairie Carneau.

Benoit, Virgil. In progress. French presence in the Red River Valley part 1: A history of the Métis to 1870.

______________. 1975. Gentilly: A French Canadian community in the Minnesota Red River Valley. In Minnesota History 1: 279-289.

Bjornson, Val. 1969. The history of Minnesota, Vol. 1. West Palm Beach, Florida: Lewis Historical Publishing Company.

Boughton, Charles E. 1976. The pioneer days of Red Lake Falls. In A History of Red Lake Falls. The Red Lake Bicentennial Committee. Dallas: Taylor Publishing Company. 109-111.

Census of the Population. 1950. Vol. II. Characteristics of the population. Part 23, Minnesota. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, U.S. G.P.O., 1952.

Census of the Population. 1980. Vol. 1. Characteristics of the population. Chapter C, General social and economic characteristics. Part 25, Minnesota. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, U.S. G.P.O., 1983.

Census of the Population. 1980. Vol. 1. Characteristics of the population. Chapter B, General population characteristics. Part 25, Minnesota. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census, U.S. G.P.O., 1982.

Compendium of the Eleventh Census. 1890. Part 1, Population. Washington, D.C.: Department of the Interior, Census Office, 1892.

Crawford, John C. 1976. Michif: A new language. In North Dakota English, 1. Grand Forks, ND: Washburn Printing Center. 3-10.

Dil, Anwar S. and Timothy Curry. 1981. Language shift in the United States: some demographic clues. In Language Diversity and Language Contact. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 158-172.

Dorrance, Ward Allison. 1935. The survival of French in the old district of Ste. Genevieve. In University of Missouri Studies, Vol. 10, No. 2. Columbia, MO: Univ. of Missouri.

Ervin-Tripp, Susan M. 1973. Identification and bilingualism. In Language Acquisition and Communicative Choice. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 1-14.

Ferguson, A. Charles, and Shirley Brice Heath. 1981. Language in the USA. Cambridge, England: Cambridge Univ. Press.

Fishman, Joshua A. 1966. Language loyalty in the United States. The Hague: Mouton.

________________. 1972a. Sociolinguistics: A brief introduction. Rowley, MA: Newbury House Publishers.

________________. 1972b. The sociology of language. In Language in Sociocultural Change. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 1-15.

________________. 1972c. Problems and prospects of the sociology of language. In Language and Sociocultural Change. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 268-285.

________________. 1972d. Societal bilingualism. In Language in Sociocultural Change. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 135-152.

Fourteenth Census of the United States taken in the year 1920. Vol. III. Composition and characteristics of the population by states. Washington, D.C.: Department of Commerce, U.S. G.P.O., 1922.

Gumperz, John J. 1972. Types of linguistic communities. In Readings in the Sociology of Language. Joshua A. Fishman, ed. The Hague: Mouton. 460-472.

Hall, Robert A. 1966. Pidgin and Creole languages. Ithaca: Cornell Univ. Press.

Haugen, Einar. 1972a. The ecology of language. In The Ecology of Language. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 325-339.

____________. 1972b. The stigmata of bilingualism. In The Ecology of Language. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 307-324.

____________. 1972c. Problems of bilingualism. In The Ecology of Language. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 59-78.

Hymes, D. H. 1962. The ethnography of speaking. In Anthropology and Human Behavior. T. Gladwin and W. C. Sturtevant, eds., Washington, D.C.:n. p. 13-53.

Karttunen, Frances. 1977. Finnish in America: A case study. In Sociocultural Dimensions of Language Change. Ben Blount and Mary Sander, eds. New York: Academic Press. 173-184.

Labov, William. 1972. Sociolinguistic Patterns. Philadelphia: Univ. of Philadelphia Press.

LePage, Robert. 1977. Processes of pidginization and creolization. In Pidgin and Creole Linguistics. Albert Valdeman, ed. Bloomington: Indiana Univ. Press. 222-258.

Lieberson, Stanley. 1981. Procedures for improving sociolinguistic surveys of language maintenance and language shift. In Language Diversity and Language Contact. Anwar S. Dil, ed. Stanford: Stanford Univ. Press. 263-280.

Locke, William N. 1949. Pronunciation of the French spoken at Brunswick, Maine. Greensboro, N.C.: American Dialect Society.

McDermott, John Francis. 1971. A glossary of Mississippi Valley French, 1673-1850. St. Louis: Washington Univ.

Mackey, William F. 1968. The description of bilingualism. In Readings in the Sociology of Language. Joshua Fishman, ed. The Hague: Mouton. 554-584.

Milroy, Lesley. 1980. Language and social networks. Baltimore: Univ. Press.

Red Lake Bicentennial Committee. 1976. A history of Red Lake County. Dallas, TX: Taylor Publishing Co.

Sixteenth Census of the United States. 1940. Population. Vol. II. Characteristics of the population. Part 7, Reports by States, Minnesota-New Mexico. Washington, D.C.: U.S. Department of Commerce, Bureau of the Census. U.S. G.P.O., 1961.

Thirteenth Census of the United States. 1910. Vol. II, Population 1910: Reports by states with statistics for counties, cities and other civil divisions. Alabama-Montana. Washington, D.C.: Department of Commerce, Bureau of the Census. U.S. G.P.O., 1913.

Thomas, Rosemary Hyde, and Ronald Thomas. 1981. It's good to tell you: French folktales from Missouri. Columbia, Missouri: Univ. of Missouri Press.

Tisch, Joseph LeSage. 1959. French in Louisana. New Orleans: A. F. Laborde and Sons.

Wolfram, Walt, and Ralph Fasold. 1974. The study of social dialects in American English. Englewood Cliffs, N.J.: Prentice Hall.